

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:26:37 ; Search time 19.49 seconds
(without alignments)
2102.715 Million cell updates/sec

Title: US-09-271-584A-2
Perfect score: 2755
Sequence: 1 MLDLSVSKLPSTSDHASV.....FVPVPGSPTRPNPDLPSKA 538
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2136	77.5	457	2 T01804	Na+/H+-exchanging
2	583.5	21.2	703	2 T26529	hypothetical prote
3	564	20.5	633	2 S69734	hypothetical prote
4	544	19.7	569	2 T37706	probable sodium/hy
5	531	19.3	629	2 T22848	hypothetical prote
6	475	17.2	832	2 A40205	Na+/H+-exchanging
7	467	17.0	831	2 B40204	Na+/H+-exchanging
8	455.5	16.5	813	2 A46748	Na+/H+-exchanging
9	454.5	16.5	809	2 A46747	Na+/H+-exchanging
10	443	16.1	815	2 S17487	Na+/H+-exchanging
11	437	15.9	816	2 S16328	Na+/H+-exchanging
12	436	15.8	818	2 A48858	Na+/H+-exchanging
13	434.5	15.8	698	2 A57644	Na+/H+-exchanging
14	431	15.6	759	2 A46188	CAMP-activated Na+
15	430.5	15.6	717	2 C40204	Na+/H+-exchanging
16	424	15.4	478	2 T18746	Na+/H+-exchanging
17	423.5	15.4	820	2 A40204	Na+/H+-exchanging
18	420	15.2	822	2 S30198	Na+/H+-exchanging
19	406	14.7	798	2 T23539	hypothetical prote
20	391.5	14.2	634	2 T33528	hypothetical prote
21	385.5	14.0	651	2 T31869	hypothetical prote
22	381	13.8	375	2 B40205	Na+/H+-exchanging
23	366.5	13.3	660	2 T28016	hypothetical prote
24	327.5	11.9	609	2 S30910	Na+/H+-exchanging
25	311.5	11.3	494	1 F69355	Na+/H+ antiporter
26	289	10.5	684	2 T16072	hypothetical prote
27	279.5	10.1	531	2 D96827	protein F20B17.4 [
28	278.5	10.1	527	1 S75063	Na+/H+-exchanging
29	254.5	9.2	575	2 D96585	hypothetical prote

ALIGNMENTS

RESULT 1

T01804
Na+/H+-exchanging protein 3 homolog A_TM021B04.4 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 22-Oct-1999
C:Accession: T01804
R:Dante, M.; Wamsley, P.; Gibson, A.
submitted to the EMBL Data Library, June 1997
A:Description: The sequence of A. thaliana TM021B04.
A:Reference number: 214440
A:Accession: T01804
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-457 <DAN>
A:Cross-references: EMBL:AF007271; NID:g2191181; PID:g2191184; GSPDB:GN00063; ATSP:A-
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:A_TM021B04.4
A:Map position: 5
A:Introns: 55/3; 95/2; 128/1; 148/1; 163/3; 240/3; 256/2; 286/3; 320/2; 400/3; 421/3

Query Match	77.5%	Score	2136;	DB	2;	Length	457;
Best Local Similarity	95.6%	Pred. No.	9.8e-155;				
Matches	430;	Conservative	0;	Mismatches	0;	Indels	20;
Gaps	2;						
QY	1	MLDSLVS	KLP	SL	STSDHAS	VV	ALNLFV
Db	1	MLDSLVS	KLP	SL	STSDHAS	VV	ALNLFV
QY	61	V	TILLIS	KG	SK	SH	LLV
Db	61	V	TILLIS	KG	SK	SH	LLV
QY	121	S	CTIIS	L	G	V	T
Db	121	S	CTIIS	L	G	V	T
QY	181	G	V	N	D	A	T
Db	181	G	V	N	D	A	T
QY	229	T	G	L	S	A	Y
Db	229	T	G	L	S	A	Y
QY	281	I	V	M	S	H	T
Db	281	I	V	M	S	H	T
QY	301	I	V	M	S	H	T
Db	301	I	V	M	S	H	T
QY	341	V	S	S	I	L	M
Db	341	V	S	S	I	L	M

probable sodium/hy
sodium/proton exch
probable Na+/H+ an
probable Na+/H+ an
conserved hypothet
Na+/H+-exchanging
hypothetical prote
probable Na+/H+ an
Na+/H+ antiporter
probable Na+/H+ an
hypothetical 60.5
probable sodium/pr
probable Na+/H+ an
probable integral
probable Na+/H+ an
Na+/H+ antiporter

Db 361 VSSILMLGVVGRAAFVPLSFLSNLAKKQSEKINFNQVIVWWSGLMRGAVSMALAYN 420

QY 401 KETRAHGTDRGNAMITTSITVCLFSTVV 430

Db 421 KETRAHGTDRGNAMITTSITVCLFSTVV 450

RESULT 2

T26529 hypothetical protein Y18D10A.6 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26529

R:Haris, B.

submitted to the EMBL Data Library, December 1998

A:Reference number: T26529

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-703 <WIL>

A:Cross-references: EMBL:AL034393; PIDN:CAA22320.1; CESP:Y18D10A.6

A:Experimental source: clone Y18D10A

C:Genetics:

A:Gene: CESP:Y18D10A.6

A:Introns: 23/3; 56/1; 103/2; 246/3; 352/3; 492/2; 578/3

Query Match 21.2%; Score 583.5; DB 2; Length 703;

Best Local Similarity 34.6%; Pred. No. 1.4e-36;

Matches 150; Conservative 79; Mismatches 160; Indels 45; Gaps 12;

QY 25 LFVALLCACIVLGHLEENRMNESITALLIGLGTGVTTLLISKSKSHLLVSEDLFFI 84

Db 138 LFVIMATLVVHMLIVSKIHMPESLAIVAGLALIG-SILSVSRDSEIEALSPDVFLL 196

QY 85 YLLPPIIFNAGFQVKKQKOFFRNFYTIMLFGAVGTIISLGTQFFKLDIGFDLG 144

Db 197 VLLPPIIFNAGFQVKKQKOFFRNFYTIMLFGAVGTIISLGTQFFKLDIGFDLG 256

QY 145 DYLAIGAIFAATDSVCTLOVNODET-PLLYSLVFGGVDNATSVVV-----F 192

Db 257 ECFAPAAISADVDPGLTAFQAVKVESLYMLVFGESMLNDASVILVATLAKHAKPSP 316

QY 193 NAIQSFOLTHLHNEAFLHGNFLYFLFLLSTLGAATGLISAYIKKLYFGRHSTDREVA 252

Db 317 NSLPASEII---TSAP---VTTEFFFSACLGVGIGLLSALLKHYDL-RKTPSLEPA 368

QY 253 LMMLMAYLSYMAELFDLSGILTVFCGIVMSHYTHWNTVTSRTTKHTFAILSLAET 312

Db 369 LLLIFSYIPYGAELDLSGILTVFCGIVMSHYTHWNTVTSRTTKHTFAILSLAET 428

QY 313 FIFLYVGMDS---ALDIDKWRVSVDTPGTSAIVSSILMGLVMVGRAAFPVPLSFLSNLAK 368

Db 429 SFYFAYIGMAFFITILNFPW-----LIFWSVV--LCILGRACNVFPLAYLVNQCR 476

QY 369 KNOSEKINFNQVIVWWSGLMRGAVSMALAYNKFTRAGHTDVRGNAMITTSITVCLFST 428

Db 477 KD--VQISMKQIILWIFSG-MRGAVCFALVLYM-----DLQEKKSILLITVFLILFTT 528

QY 429 VVFGMLTKPLISYL 442

Db 529 IFLGGSALPFIISFI 542

RESULT 3

S69734

hypothetical protein YDR456w - yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*

C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001

C:Accession: S69734

R:Dieckrich, F.S.

submitted to the EMBL Data Library, August 1995

A:Description: The sequence of *S. cerevisiae* lambda 3641 and cosmids 9461, 9831, and 9411

A:Reference number: S69555

A:Accession: S69734

A:Molecule type: DNA

A:Residues: 1-633 <DIE>

A:Cross-references: EMBL:U33007; NID:9927685; PID:9927695; GSPDB:GN00004; MIPS:YDR456

C:Genetics:

A:Gene: MIPS:YDR456w

A:Map position: 4R

C:Superfamily: hypothetical protein ywgp

Query Match 20.5%; Score 564; DB 2; Length 633;

Best Local Similarity 26.8%; Pred. No. 3.8e-35;

Matches 164; Conservative 121; Mismatches 200; Indels 128; Gaps 21;

QY 14 TSDHASVVALNLFVALLCACIVLGHLEEN--RWNNESITALLIGLGTGVTTLLISKSKS 71

Db 55 TEEMFSSWALFTMLLLISALWSSYYLTQKRIRAVHETVLSIFYGMVIGLIIRMSPGHYI 114

QY 72 SHLVFESDLFFIYLLPPIIFNAGFQVKKQKOFFRNFYTIMLFGAVGTIISCTIISLGTQ 131

Db 115 QDVTVTNSSFYFNVLLPPIILNSGVNELQNVNFFNNMLSILIFAIPTGTFISAVVIGI-ILY 173

QY 132 FRKKDICTFDL--GDYLAIGAIFAATDSVCTLOVNODET-PLLYSLVFGGVDNATSD 188

Db 174 IWTFLGESIDISFADAMSVGATLSATDPVTIISIFNAYKVDPKLYIIIFGESLNDALS 233

QY 189 VVFNIAIQSFDTLHNEAFLHGNFLYFLFLLSTLGAATGLISAYIKKLYFGRHSTD 248

Db 234 IVNFETCORFHQOPATFSVFEAGLFLMTFESVLLIGVLIGLIVALLKHTHIRRY-PQ 292

QY 249 REVALMMLMAYLSYMAELFDLSGILTVFCGIVMSHYTHWNTVTSRTTKHTFAILSF 308

Db 293 IESCLILLIAYESYFSGCHMSGIVSLLECGITIKHYAYNNMSRSQITIKYIQLLAR 352

QY 309 LATFTFIFYGMDALDIDKWRVSVDTPGTS-----TAVSSILMGLVMVGRAAFPVFP 359

Db 353 LSENFIFYLGLLELF-----TEVELYKPLLIIVAAI---SICVARWCAPVP 396

QY 360 LS-----FLSNLAKKQNS--EKINFMQVIVWWSGLMRGAVSMALAYN---- 400

Db 397 LSQFVWNIYRVKTIKNSGITGENISVDEIPYNYQMTFWAGL-RGAVGVVALGIQGE 455

QY 401 -KTRAGHTDVRGNAMITTSITVCLFSTVFGMLTKPLISYLLPHONATTSMLSDDN- 458

Db 456 YKFT-----LLATVLVVVLTIVIFGCTTAGMLEV----NIKTGCSIEDTS 499

QY 459 -----PKSIHI-----PLLDQDSFIEFSGHNHVPDPDSTIRGELT 492

Db 500 DDEFDIEAPRALNLLNGSSIQTDLPYSDNNSPDISIDQF-AVSSNNKLPNNISTTGNT 558

QY 493 -----RPTRTV-----HYWRQFDDSPFMRVFGGGRGVFPV 523

Db 559 FGGLNETENTSPNPARSSMDKRNLRDLKLTIFNSDSQWQNFQDEQVLKPVFLD-NVSPSL 617

QY 524 PGSPTRNPPDLS 536

Db 618 QDSATQ-SPADES 629

RESULT 4

T37706

probable sodium/hydrogen exchanger - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: T37706

R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1997

A:Reference number: Z17138

A:Accession: T37706

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-569 <MUR>

QY 76 VFSEDLFFIYLLPPIIFNAGFQVKKOFFERFVTIMLFGAVGTIISC-----TI 124
||| :|| |:|||||||: :||: ||| :|| ||| :|| :||

QY	20	VVALNLFVALLCACIV-LGHLLLENRMNSESITALLIGLIGTGVTTILLISKGKSHLLVP-	77
Db	56	IIALWLVAFLAKIVFHLSH--KVTSVVPESALLIVGLVGLVGIVL-----AADHIASET	108
QY	78	-SDSLFFIYLLPPIIFNAGFOVKKQFNRFTVIMLFVGAVGTIIISTIIISLGVTFQFKKL	136
Db	109	LTPTEVFYLLPPIVDAGYFMPNRLFEFNSLGSIIILYAVGVVWNAITGLSLXGVFLSG	168
QY	137	DIGTFDLG--DYLAIGATFAATDSVCTQLVNO--DETPLYLSLVFEGGVVNDATSVVFN	193
Db	169	IMGELKIGLLDLFEGSLIAANDPVAVLFESVHNVEVLFTIVFEGSILLANDAVTVVLYN	228
QY	194	AIGSFDLTHLNEAAFHLLGNLFVFLFLLSLLCAGATGLISAYVKKL--YFCGRHSTDEVA	252
Db	229	VFQSFVTLGGDKVTGDCVKIGISVFVVS--LGGTLGVVYFAFLSLVTRFKHVRVLEPG	287
QY	253	LMLMAYLSYMLAELFDLSGLITVFVFCGIVMSHYTHVNTSESSRITTKHTPATISFAET	312
Db	288	FVPIIISYLTSEMLSSLSLITATFCGICOKYKVRANISEQATTVRYTWKMLASGAET	347

RESULT 9		A:Status: preliminary; translated from GB/EMBL/DBD	
A46747		A:Molecule type: mRNA	
Na+/H+-exchanging protein NHE-2 - rabbit		A:Residues: 1-815 <FLI>	
C:Species: Oryctolagus cuniculus (domestic rabbit)		A:Cross-references: GB:S68616; NID:g544775; PIDN:AAC60606.1; PID:g544776	
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997		R:Sardet, C.; Franchi, A.; Pouyssegur, J.	
C:Accession: A46747		Cell 56, 271-280, 1989	
R:Tse, C.M.; Levine, S.A.; Yun, C.H.; Montrose, M.H.; Little, P.J.; Pouyssegur, J.; Dono		A:Title: Molecular cloning, primary structure, and expression of the human growth fac	
J. Biol. Chem. 268, 11917-11924, 1993		A:Reference number: A31311; MUID:89106219	
A:Title: Cloning and expression of a rabbit cDNA encoding a serum-activated ethylisoprop		A:Accession: A31311	
A:Reference number: A46747; MUID:93280159		A:Molecule type: mRNA	
A:Accession: A46747		A:Residues: 1-814/'SNARASQRSLPHRLFHQRCWGLPFLTRIGPAPPPTAWQLGPPQPHQSSSPASREASSPP	
A:Status: preliminary		C:Genetics:	
A:Molecule type: nucleic acid		A:Gene: GDB:SLC9A1; APNH; NHE1	
A:Residues: 1-809 <TSE>		A:Cross-references: GDB:119683; OMIM:107310	
A:Experimental source: ileal villus cells		A:Map position: lp36.1-lp35	
A>Note: sequence extracted from NCBI backbone (NCBIN:133350, NCBIP:133351)		C:Keywords: glycoprotein; transmembrane protein	
		F:75,370,410/Binding site: carbohydrate (Asn) (covalent) #status predicted	
		Query Match 16.1%; Score 443; DB 2; Length 815;	
		Best Local Similarity 26.0%; Pred. No. 8.1e-26;	
		Matches 151; Conservative 99; Mismatches 205; Indels 126; Gaps 22;	
QY	7 SKLPSTSDHASV---VALNLEFVALLCACIVLGHLLLEE-NRWNNESITALLIGLGTGVTT	QY	23 LNLFVALLCACIVLGHLLLEE-NRWNNESITALLIGLGTGVTTLLISKSSHLLVFSDDL 81
Db	66 SRLP-VFTLDYPHVQIPFEITLILLASLAKIGFHLVHKLPTIVPESCLLIMVGLLGGI 124	Db	105 ISLWILLACLAKMKIGFHVPTISSIVPESCLLIVVGLLVGGLI---KGVGETPPPLQSDV 160
QY	63 ILLISKSSHLLVFSDDLFIYLLPIIPNAGQVKKKQFFNFVTIMLFGAVGTIIC 122	QY	82 FIYLLPIIPNAGQVKKKQFFNFVTIMLFGAVGTI-----ISCTIISLGVTQFF 133
Db	125 IFGVDE-KSPAM--KTDVPFLYLLPVIDLDAGYFMPTRPFENLGTIFEWYAVVGLTWS 181	Db	161 FFLFLLPIILDAGYFLPQRQFTENLGTILIFAVVGLTWNAAFLLGGLVAVCLVGGQE-- 218
QY	123 TIISLGVTOFFKLDIGTDLGDY----LAIGALFAATDSVCTLOVL-NODETPLLYSL 176	QY	134 KKLDIGTDLGDYLAIGALFAATDSVCTLOVLNODE-TPLLYSLVFGCVNDATSVVVF 192
Db	182 --IGIGVS-LFGICQIEAGLSOITLQNLFLGSLISAVDPVAVLAVFENIHVNEQLYIL 238	Db	219 ----INNIGLDNLFGSIISAVDPVAVLAVFEEIHNEHLHILVFGESLLNDVAVTVVLY 274
QY	177 VFGEGVNDATSVVFNIAQSF-DLTHLNHEAAFLHLLGNFLYLLSTLIGAAATGLISAY 235	QY	193 NAIQSFDLTHLNHEAAFLHLLGNFLYLLS---TLGAAATGLISAYVIKKLYFGRHSTD 249
Db	239 VFGESLNDATVVVLYNLFKSFQOMKTIEDYFAGIANEFVVVGIGVLLIGIFLGFIAF 298	Db	275 HLFEF--ANYEHVGVIVDIFLGFSLFFVVLGGVLVGVVGVIAFTSR---FTSHIRVI 329
QY	236 VIKKLYFGRHSTREVALMMAYLSYMLAEFLDLSGILTVFFCGIVMSHYTHWNVTSS 295	QY	250 EVALMMAYLSYMLAEFLDLSGILTVFFCGIVMSHYTHWNVTSSRTTKHTFATLSFL 309
Db	299 TTR---PTHNIRVIEPLFVLYSLYITAEFHLGSIIMAITACAMTKNRYEENVSOKS 355	Db	330 EPLFVLYSYMAYLSAEFLHLSGIMALIASGVVMPYPYVEANISHKSHTTIKYFLKMWSV 389
QY	296 RITTKHFTATLSFAETFIYLYGMDALDID---KRSVSDTPTGTSIAVSSILLMGLVMVG 352	QY	310 ATEFTFLVYGMDALDIDKWSVSDTPTG-----STAVSSILLMGLVMVGRAAFPVPLS 364
Db	356 YTIKYPKMLSSVSETLIFIFMGVSTVGKNHEWNAFVCFI-----LAFCLIW 404	Db	390 SETLIFIFLG-----VSTVAGSHHNNWTIVISTLL--FCLIARVILGVGLTWFI 436
QY	353 RAAPVPLSFLSLNLAQSEKINFNMQVIVMWSGLMRGAVSMALAY----NKFTRAGHT 408	QY	365 NLAKNQSEKINFNMQVIVMWSGLMRGAVSMALAY----NKFTRAGHTDVRGNAIMITST 420
Db	405 RALGVFVLTIVNWR---TIPLTFKQDFIIAYGGL-RGAICFALVFLPAAVFPF--- 456	Db	437 N---KFRIVKLTPKDQFIAYGGL-RGAIAFSLGYLLDKKHPPMCD-----LFLTAI 484
QY	409 DVRGNATMITSTIVCLFSTVFGMLTKPLISYL-LPHQNATTSMLSDNTPKSIHIPL 467	QY	421 ITVCLFSTVFGMLTKPLISYLLPHQNATTSMLSDNTPKSIHIPLLDODSFIEPSSGHNH 480
Db	457 ----KKLFITAAIVIFFTFVILGITRPLVLEFVLKRSNKKQQAQVSEE-----IHCREF 507	Db	485 ITVIEFTVFGQMTIRPLVDLAVKKQKQETKR---SINEEHTQFLD-----HL 530
QY	468 DQ-DSFTEPSGNINVRPDSIRGFLTRPTVHYWR---QFDDSEMR 511	QY	481 VPRPDSIRGFLTRPTVHYWR---QFDDSEMRPVE----- 514
Db	508 DHVKTGIEDVCGH-----WGHNFRWDFKFKFDDXYLR 539	Db	531 LTGIEDICGHY-----HHHMKDKLNRFNKYVKKLLIAGERSKEPOLIAFYHKMEMKQ 584
RESULT 10		QY	515 -----GGRGFVPFV-----PGS-PTERNPPDLSK 537
I57487		Db	585 ATELVEGSGMGKIPSAVSTVSMQNIHPKSLPSEITLPALSK 625
Na+/H+-exchanging protein, amiloride-sensitive - human		RESULT 11	
N:Alternate names: Na+/H+ antiporter; NHE-1		SI6328	
C:Species: Homo sapiens (man)		Na+/H+-exchanging protein - rabbit	
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Oct-1999		N:Alternate names: Na+/H+ antiporter; Na+/H+ exchanger; pH regulatory protein	
C:Accession: I57487; A31311		C:Species: Oryctolagus cuniculus (domestic rabbit)	
R:Fliegel, L.; Dyck, J.R.; Wang, H.; Fong, C.; Haworth, R.S.		C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 01-Dec-2000	
Mol. Cell. Biochem. 125, 137-143, 1993		C:Accession: SI6328; SI3926; S30602	
A:Title: Cloning and analysis of the human myocardial Na+/H+ exchanger.		R:Tse, C.M.; Ma, A.I.; Yang, V.W.; Watson, A.J.M.; Levine, S.; Montrose, M.H.; Potter	
A:Reference number: I57487; MUID:94111706		EMBO J. 10, 1957-1967, 1991	
A:Accession: I57487			

Db 485 ITVFTFVQGMTRIRPLVDLAVKKQETKR-----SINEIHTQFLD-----HL 530
Qy 481 VPRPDSIRGLRPTRTVHYWYR-----QFDDSMRPVF----- 514
Db 531 LTGIEDICHYG-----HHWKDKLNRNKKYVKKKLIAGERSKEPQOLIAFYHKMEMKO 584
Qy 515 -----GGRGVFPVPGS-----PTERNPPDLCK 537
Db 585 AIELVESGGMKIPSAVSTVSMQNIHPKTLERILPALS 625
RESULT 13
A57644
Na+/H+-exchanging protein 2 - human
C:Species: Homo sapiens (man)
C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 28-May-1999
C:Accession: A57644
R:Ghoshan, F.K.; Knobel, S.M.; Summar, M.
Genomics 30, 25-30, 1995
A:Title: Molecular cloning, sequencing, chromosomal localization, and tissue distribution of the human Na+/H+-exchanging protein 2 (NHE2) gene.
A:Reference number: A57644; MUID:96129297
A:Accession: A57644
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-698 <GHI>
A:Cross-references: GB:S81591; MID:g1478393; PIDN:AAB36180.1; PID:g1478394
C:Genetics:
A:Gene: GDB:SLC9A2
A:Cross-references: GDB:132596; OMIM:600530
A:Map position: 2q11.2-q11.2
C:Keywords: transmembrane protein

Query Match 15.8%; Score 434.5; DB 2; Length 698;
Best Local Similarity 28.5%; Pred. No. 2.9e-25;
Matches 137; Conservative 84; Mismatches 182; Indels 77; Gaps 20;
Qy 54 LIGLGTGVTLLISKGSLLHVFSEDLFFIYLLPPIIFNAGQVKKKQFFRNFVIMLF 113
Db 1 MVGLLGGIIFGVDE-KSPAM-KTDVEFLYLLPPIVDAGVFMPTRPFFENLGTIFW 57
Qy 114 GAVGTIISLISLVG-TQFFKLDIGTDLGDI-----LAIGAFAATDSVCTLOVL-N 166
Db 58 AVGTIWN-----SIGISLFGICQIEAFGLSDITLQNLFGSLISAVDPVAVLAVFN 113
Qy 167 QDETPLYLSLVFGGVNDATSVVFNIAQSF-DLTHLNHEAFAHLLGNFLYLLFLLSTLL 225
Db 114 IHVNEQLYTLVFGESLLNDVAVLYLNLFKSCQMKTIQTVDFAGIANFVVGIGVLI 173
Qy 226 GAATGLISAVVKKLYFGHSTDEVALMMLMAYLSYMLAEFLDLGILTVPFCGIVMSH 285
Db 174 GILLGFIAFTTR---FTHNIRVIEPLFVLYSLYSLTITAEHFLSGIMAITACAMTKN 230
Qy 286 YTHNVTESRITRHTFATLSFAETIFLYVGMADLID---KWRKSVSDTPGTSIAVS 342
Db 231 YVEENVSKSYTTIKYFMKLSVSETLIFFMGVSTVGKNHEWNAFVCF----- 282
Qy 343 SILMGLVMVGRAAFPVPLSLNKKQSEKINFNQVWVWGLMRGAVSMALAY--- 399
Db 283 ---LATCLWRALGVFLVQVFNFR---TIPLFKQDFIAYGGL-RGAICFALVFLP 335
Qy 400 -NKEFRAGHTDVRGNAMITSTITVCLFSTVVFGLMKPLISYL-LPHQNATTSMLSDN 457
Db 336 ATVFPFR-----KKLFTAAIAVIVFTVPLILGTTIRPLVEFLDKVRKSKQOAVSEE- 386
Qy 458 TPKSIH-IPLLDQ-DSFIEPSGNHNVPRPDSIRGLRPTRTVHYWYR-----QFDDSMR 511
Db 387 ----IHWCRFDHVKTGIEDVCGH-----WGHNFWRDKFKFDKYL 425

RESULT 14

A46188

CAMP-activated Na+/H+-exchanging protein betaNHE - rainbow trout

C:Species: Oncorhynchus mykiss (rainbow trout)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997
C:Accession: A46188
R:Borgese, F.; Sardet, C.; Cappadoro, M.; Pouyssegur, J.; Motaïs, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 6765-6769, 1992
A:Title: Cloning and expression of a CAMP-activated Na+/H+ exchanger: evidence that it is a member of the Na+/H+ exchanger gene family.
A:Reference number: A46188; MUID:92337712
A:Accession: A46188
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-759 <BOR>
A:Experimental source: cephalic kidney hematopoietic tissues
A:Note: sequence extracted from NCBI backbone (NCBIP:113525)

Query Match 15.6%; Score 431; DB 2; Length 759;
Best Local Similarity 27.3%; Pred. No. 6e-25;
Matches 145; Conservative 102; Mismatches 183; Indels 102; Gaps 22;
Qy 21 VALNLEVALCACIVLG-HLLEE-NRMNSESITALLIGLGTGTILLISKSSHLVFS 78
Db 73 IALWILLALL---MKLGFHLIPRLSAVVPESCLLVVGLLVGGLIKVIGBEPD---VLD 125
Qy 79 EDLFFIYLLPPIIFNAGQVKKKQFFRNFVIMLFAGVGTIIS-----CTIISL 127
Db 126 SQLEFFCLLPPIILDAGYFLRPTENVTENVTILFVAVIGTLWNAFFMGGLLYALCOLESV 185
Qy 128 GVTQFFKLDIGTDLGDI-LAIGAFAATDSVCTLOVLNQE-TPLIYSLVFGGVVND 186
Db 186 G-----LSGVDLLACLLFGSIVSAVDPVAVLAVFEEHINELVHILVFGESLLND 236
Qy 187 TSVVVENAIQSFED---LTHLNHEAFAHLLGNFLYLLSTLLGAATGLISAVIKKLY 242
Db 237 VTVLVNLNLEEFKSVGTVTLD---VFLGVVCFVFLSGVGLVGAIFYLAFTSR---F 290
Qy 243 GRHSTDREVALMMLMAYLSYMLAEFLDLGILTVPFCGIVMSHYTHNVTESRITRHT 302
Db 291 TSHTRVIEPLFVLYSYMAYLSSEMFHLSGIMALIAGVVMRYVEANISHKSYTTIKYF 350
Qy 303 FATLSFLAETIFLYVGMADLIDKWRKSVSDTPGT-----SIAVSSILMGLVMVGRAAV 357
Db 351 LKMWSSVSETLIFFLG-----VSIVAGPHANNFTVITVI--LCLVSRVLGV 397
Qy 358 FPLGSFSLNKKQSEKINFNQVWVWGLMRGAVSMALAYNKKFRAGTIDVRGNA--- 414
Db 398 IGLTFIN---KFRIVKLTQKQDFIAYGGL-RGAIAFSLGY-----LLSNHOM 443
Qy 415 --LMTSTITVCLFSTVVFGLMTRPLISYLPHQNATTSMLSDNTPKSIHILPDDQSF 472
Db 444 RNLFALITVITVFFVQGMTRIRPLVEALLVKKKE---SKPSINEEHTFEFLD--- 495
Qy 473 IEPGSHNVPRPDSIRGLRPTRTVHYWYRQ---FDDSEFM-RPVFGGRGF 519
Db 496 -----HLITVEGVCGHYG-----HYHWKELNRFNKTYVVRKLIAGENF 535

RESULT 15

C40204

Na+/H+-exchanging protein 4 - rat

N:Alternace names: Na+/H+ antiporter

C:Species: Rattus norvegicus (Norway rat)

C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 05-Nov-1999

C:Accession: C40204

R:Orlowski, J.; Kandasamy, R.A.; Shull, G.E.

J. Biol. Chem. 267, 9331-9339, 1992

A:Title: Molecular cloning of putative members of the Na/H exchanger gene family. cDNAs for the rat Na/H exchanger protein 4 (NHE4) and the rat Na/H exchanger protein 5 (NHE5) were isolated and sequenced. The deduced amino acid sequences of these proteins show that they are members of the Na/H exchanger gene family.

A:Reference number: A40204; MUID:92250539

A:Accession: C40204

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-717 <ORL>

A:Cross-references: GB:M85301; MID:g205708; PIDN:AAA41703.1; PID:g205709

C:Keywords: transmembrane protein

```
Query Match      15.6%; Score 430.5; DB 2; Length 717;
Best Local Similarity 26.2%; Pred. No. 6.1e-25;
Matches 137; Conservative 102; Mismatches 189; Indels 95; Gaps 21;

QY 36 LCHLLEENRWNESITALLIGTGVTILLISKGSSHLVFSDDLFFIYLLPPIFNAG 95
Db 93 LPHL-----MPESCLLIIVGALVG-SIIFGTHKSP--VMDSSIYFLYLLPPIVLESG 143
QY 96 FOVKKKOFFERFVIMLFAGAVGTHISCTIISLGVTOFF----KKLDICTFDLGDYLAIGA 151
Db 144 YFMPTRPFEPFENIGSILWAGLALINA--FGICLSLYFICQAKAFGLGDINLLQNLLFGS 201
QY 152 IFAATDSVCTQLVNLQDE-TPLLYSLVFGEGVVDNATSVVFNAIQSFDLTHL--NHERA 208
Db 202 LISAVDPVAVLAFEEARVNEQLYMMIFGEALLNDGISVLYNILLIAFTKHKFEDEIADV 261
QY 209 FHLIG--NFLYLLSTLLGAATGLISAVIKKLYFGRHSTDRVALMMLMAYLSYMLAE 266
Db 262 DILAGCAREFIVGGGVFFGIIFFGIFSAFTR---FTQNISAIPELIVFMFSYLSVIAAE 318
QY 267 LFDLSGILTVFFCGIVMSHYTHNVTESSRIITKHTFATLSFLAETFIPLYVGMALDID 326
Db 319 TLYLSGILAITACAVTMKKYVEENVSTSYTTIKYFMKMLSSVSETLIFIFMGVSTVGKN 378
QY 327 ---KWRVSVDTPGTSIAVSSILMGLVMVGRAAFEPLSFLSNLAKKNOSEKINFN--QV 381
Db 379 HEWNMAFVCF-----LAFQIWRALISVETLFYS-----NQRTFFPSINDQL 422
QY 382 VLVWSGLMRGAVSMALAY-----NKFTAGHTDVRGNAMITSTIVCLFSTVVFGLMTRP 437
Db 423 IIFYSG-VRGAGSFLAFLPLFLFPR-----KKLFVATILVVTYFTVFFQGITIGP 473
QY 438 LISYLLPHONATTSMLSDDNTPKSIHPLLDQDSFIEPSGMHNVPRDPSIRGELTRPT 497
Db 474 LVRYLDVRKTNKESINEE-----LHRLMD-----HLKAGIEDVCG-----QW 512
QY 498 VHYI---WROFDDSFMRPVFGGGEVPEVPGSPTERNPPDLS 536
Db 513 SHYQVRDKFKKFDHRYLRKIL-----IRRNQPKSS 542
```

Search completed: November 17, 2001, 13:28:25
Job time: 108 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:27:12 ; Search time 13.15 seconds
(without alignments)
1401.479 Million cell updates/sec

Title: US-09-271-584A-2
Perfect score: 2755
Sequence: 1 MDSLVKPLSLSTDSHVS.....FVPFVGSPTERNPPDLSKA 538

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	581	21.1	669	1 NAH6_HUMAN	Q92581 homo sapien
2	564	20.5	633	1 NAH2_YEAST	Q04121 saccharomyc
3	475	17.2	832	1 NAH3_RABIT	P26432 oryctolagus
4	467	17.0	831	1 NAH3_RAT	P26433 rattus norv
5	466.5	16.9	896	1 NAH5_HUMAN	Q14940 homo sapien
6	466.5	16.9	898	1 NAH5_RAT	Q920x2 rattus norv
7	462	16.8	834	1 NAH3_HUMAN	P48764 homo sapien
8	459	16.7	839	1 NAH3_DIDMA	Q28362 didelphis m
9	455.5	16.5	813	1 NAH2_RAT	P48763 rattus norv
10	454.5	16.5	809	1 NAH2_RABIT	P50482 oryctolagus
11	452.5	16.4	812	1 NAH2_HUMAN	Q9ubyo0 homo sapien
12	443	16.1	815	1 NAH1_HUMAN	P19634 homo sapien
13	437	15.9	816	1 NAH1_RABIT	P23791 oryctolagus
14	436	15.8	818	1 NAH1_PIG	P48762 sus scrofa
15	431	15.6	759	1 NAH6_ONCMY	Q01345 oncorhynch
16	431	15.6	820	1 NAH1_RAT	P26431 rattus norv
17	430.5	15.6	717	1 NAH4_RAT	P26434 rattus norv
18	429	15.6	817	1 NAH1_BOVIN	Q28036 bos taurus
19	424	15.4	820	1 NAH1_MOUSE	Q61165 mus musculu
20	420	15.2	822	1 NAH1_CRIGR	P48761 cricetus
21	327.5	11.9	609	1 NAH1_CABEL	P35449 caenorhabdi
22	202	7.3	549	1 YJCE_ECOLI	P32703 escherichia
23	190	6.9	542	1 YMB7_MYCTU	Q50678 mycobacteri
24	160.5	5.8	578	1 YCGO_ECOLI	P76007 escherichia
25	145	5.3	383	1 NAGA_ENTHR	P26235 enterococcu
26	144.5	5.2	426	1 Y057_METJA	Q60362 methanococc
27	140.5	5.1	759	1 NAH2_SCHPO	Q14123 schizosacch
28	128.5	4.7	422	1 YF21_METJA	Q58916 methanococc
29	121	4.4	407	1 GLUP_HELPY	Q25788 helicobacte
30	120.5	4.4	808	1 ATK2_2YCRO	Q42701 zygosacchar
31	119.5	4.3	581	1 ATKA_ANASL	Q9f6x2 anabaena sp
32	117.5	4.3	430	1 RFBX_SALTY	P26400 salmonella
33	117	4.2	558	1 YBAL_ECOLI	P39830 escherichia

34	116.5	4.2	614	1 NUOL_BUCAI	P57262 buchnera ap
35	113.5	4.1	549	1 COX1_LEITA	P14544 leishmania
36	113.5	4.1	549	1 COX1_TRYBB	P04371 trypanosoma
37	113	4.1	407	1 GLUP_HELPY	Q92k41 helicobacte
38	113	4.1	519	1 NU4M_FODAN	P15582 podospora a
39	111.5	4.0	459	1 NU4M_RABIT	O79436 oryctolagus
40	111.5	4.0	468	1 NAH_SCHPO	P36606 schizosacch
41	111	4.0	553	1 NUOM_MYCTU	O53307 mycobacteri
42	111	4.0	815	1 AOX1_AERPE	Q9ydx6 aeropyrum p
43	110.5	4.0	459	1 NU4M_HIPAM	O92zy2 hippopotamu
44	107.5	3.9	402	1 Y718_METJA	O58128 methanococ
45	107.5	3.9	514	1 COX1_PELSU	O79672 pelomedusa

ALIGNMENTS

RESULT 1
NAH6_HUMAN
ID NAH6_HUMAN STANDARD; PRT; 669 AA.
AC Q92581;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 6 (NA(+)/H(+) EXCHANGER 6) (NHE-6).
GN SLC9A6 OR NHE6 OR KIAA0267.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98175963; PubMed=9507001;
RA Numata M., Petrecca K., Lake N., Orłowski J.;
RT "Identification of a mitochondrial Na⁺/H⁺ exchanger.";
RL J. Biol. Chem. 273:6951-6959(1998).
RN [2]
RP SEQUENCE OF 4-669 FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y.,
Chara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
analysis of cDNA clones from cell line KG-1 and brain.";
RL DNA Res. 3:321-329(1996).
CC -I- FUNCTION: ELECTRONEUTRAL EXCHANGE OF PROTONS FOR NA⁺ AND K⁺ ACROSS
THE MITOCHONDRIAL INNER MEMBRANE. CONTRIBUTES TO ORGANELLAR VOLUME
AND CALCIUM HOMEOSTASIS.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRION.
CC -I- TISSUE SPECIFICITY: UBIQUITOUS; BUT IS MOST ABUNDANT IN
MITOCHONDRIUM-RICH TISSUES SUCH AS BRAIN, SKELETAL MUSCLE, AND
HEART.
CC -I- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF030409; AAC39643.1; -;
DR EMBL; D87743; BAA13449.1; -;
DR MIM; 300231; -;
DR InterPro; IPR000676; -;
DR InterPro; IPR002090; -;
DR Pfam; PF00999; Na.H.Exchanger; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01086; NAHEXCHNGR6.
DR Transmembrane; Sodium transport; Transport; Symport; Mitochondrion.
KW TRANSMEM 28 48 POTENTIAL.
FT

FT	TRANSMEM	71	91	POTENTIAL.	
FT	TRANSMEM	103	123	POTENTIAL.	
FT	TRANSMEM	142	162	POTENTIAL.	
FT	TRANSMEM	179	199	POTENTIAL.	
FT	TRANSMEM	220	240	POTENTIAL.	
FT	TRANSMEM	246	266	POTENTIAL.	
FT	TRANSMEM	292	312	POTENTIAL.	
FT	TRANSMEM	340	360	POTENTIAL.	
FT	TRANSMEM	382	402	POTENTIAL.	
FT	TRANSMEM	404	424	POTENTIAL.	
FT	TRANSMEM	447	467	POTENTIAL.	
FT	TRANSMEM	483	503	POTENTIAL.	
SQ	SEQUENCE	669 AA;	74161 MW;	F6416596229F2639	CRC64;
Query Match 21.1%; Score 581; DB 1; Length 669;					
Best Local Similarity 30.0%; Pred. No. 6.9e-32;					
Matches 166; Conservative 102; Mismatches 173; Indels 112; Gaps 20;					
QY	23	LNLFVALLCACIVLGHLEEN--RWNESITALLIGLGTGVTL--	-----	65	
DB	74	LLIFILLTTLITLWLFHRRARFUEHGLAMIYGLLVGLVLRVGIHVPSDVNNTLSC	133		
QY	66	-ISKGSHLLVSEDLFFIYLPPIIFNAGFQVKKQFFNFVIMLFGAVGTIISCTI	124		
DB	134	EVQSSPTLLVTFDPEVFVFNILLPIIFVAGYSLKRRHFFRNLSILAYFLGTAISCFV	193		
QY	125	IS---LGVTOFFKKLD--IGTFDLGDYLAIGAFAATDSVCYLQVNLQDETPL-LYSLVF	178		
DB	194	IGSIMYCVTLMLKVTGQAGDFYFTDCLLFGAIVSATDPVTVLAIFHELOVDVELYALLF	253		
QY	179	GEGVNDATSVVVFNAI-----QSFDLTHLNEAAFLHLLGNFLYLFLLSTLLGAA	228		
DB	254	GESVLNDAVAIVLSSSIVAVQAGDNSHTFDVT-----AMFKSIGLFLGFSFAMGAA	308		
QY	229	TGLISAYVIK--KLYFGRHSTDREVALMMLMAYLSVLMALFPLDLSGLIVFFCGIWMVSHY	286		
DB	309	TGVVTVLVTFTKL---REFQLLETGLFLLMSWSTLLAEAWGFTGVVAVLFCGITOQAHY	365		
QY	287	TWNNVTESSRTYKHTATLSFLAEAFIFLYVGDMDALDIDKWSVSDTPTCTSAVSILM	346		
DB	366	TYNLSSTQHRTKQPELLNLAENFIIFISGLTFL-----TQNHVFNFTFVV	415		
QY	347	G---LVMVGRAAFVPLSFLSLAKKNOSEKINFNMQVLIWWSGLMRGAVSMALAYNKFT	403		
DB	416	GAFVAIFLGRAANIYPLSLNLNLRSS---KIGSNFOHMMFAFL-RGAMAFALAIR---	468		
QY	404	RAGTVDVRGNAMITSTITVCLSTFVVFGMLTKPLISYLLPHQNATTSMLSDONTPKSIH	463		
DB	469	---DTATYAROMFSTLLIVFTVWVFG-----GGTAMLS-----CLH	505		
QY	464	IPL---LDODSFIEPSGNHNVPRPDSIRGFLTRPTRVHYVYRQFDDSFMRPVFGGRGV	520		
DB	506	IRVGVDSDQEHGLVPENERTTKAES--AWLFR-----MWYFNHNLKPLL-----	550		
QY	521	PFVPGSPTERNPP	533		
DB	551	-----THSGPP	556		
RESULT 2					
ID	NAH2_YEAST	STANDARD;	PRT;	633 AA.	
AC	Q04121;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	MITOCHONDRIAL SODIUM/HYDROGEN EXCHANGER (MITOCHONDRIAL NA(+)/H(+)				
DE	EXCHANGER).				
GN	NAH2 OR NHX1 OR YDR456W OR D9461.40.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.				

Db 293 IESCLILLIAYESVFFSGCHMSGIVSLLFCGITLKHAYAYNRRSQITIKYIFOLLAR 352

Qy 309 LAETFIYVGMADLIDDKWRSVDPGTS-----IAVSSILMLGMVGRFAFVP 359

Db 353 LSENFIIYGLLELF-----TEVELVYKPLLIIVAAI---SICVARNCAVFP 396

Qy 360 LS-----FLSNLAKNQS--EKINFNMQVVIWMSGLMRGAYSMALAYN----400

Db 397 LSQFVNNIYRVKTIKRSMSGITGENISVPDEIPYQMTFTWAGL-RGAVGVALALGIQGE 455

Qy 401 -KTRAGHTDVRGNAIMTITIVCVLSTVVFGLTKPLISYLLPHONATTSMLSDNT- 458

Db 456 YKET-----LLATVLVVVLTVIIFGTTAGMLEVL---NKTGCISEEDTS 499

Qy 459 -----PKSIHT-----PLLDQDSFIEPSGNHNPVPRDSIRGFLT 492

Db 500 DDEFIEAPRAINLLNGSSIQDTGLGYPSSDNNSPDISIDOF-AVSSKNLNPNNISTTGGNT 558

Qy 493 -----RPRTRV-----HYWRQFDDSPMRPVFGGRGVFPV 523

Db 559 FGGLNETENTSPNPARSMDKRNRLKGLTIFNSDSQWFQNEQVLKPVFLD-NVPSPL 617

Qy 524 PGSPTERNPPDLS 536

Db 618 QDSATO-SPADFS 629

RESULT 3

NAH3_RABIT STANDARD; PRT; 832 AA.

AC R26432;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE SODIUM/HYDROGEN EXCHANGER 3 (NA(+)/H(+) EXCHANGER 3) (NHE-3).

GN SIC9A3 OR NHE3.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NEW ZEALAND WHITE; TISSUE=ileal villus, and kidney cortex;

RX MEDLINE=92250340; PubMed=1374392;

RA Tse C.-M., Brant S.R., Walker S.S., Pouysseur J., Donowitz M.;

RT "Cloning and sequencing of a rabbit cDNA encoding an intestinal and

RT kidney-specific Na+/H+ exchanger isoform (NHE-3).";

RL J. Biol. Chem. 267:9340-9346(1992).

CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED

CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL

CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD

CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL

CC TRANSDUCTION.

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -!- TISSUE SPECIFICITY: INTESTINAL AND KIDNEY SPECIFIC. MOST ABUNDANT

CC IN KIDNEY CORTEX, FOLLOWED EQUALLY BY ILEUM AND ASCENDING COLON,

CC THEN KIDNEY MEDULLA AND JEJUNUM. IS ABSENT FROM DUODENUM AND

CC DESCENDING COLON.

CC -!- PTM: PHOSPHORYLATED (POSSIBLE).

CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.

CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC

CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M87007; AAA31420.1; -

DR PIR; A40205; A40205.

DR InterPro: IPR000676; -

DR InterPro: IPR001118; -

DR Pfam: PF00999; Na_H_Exchange; 1.

DR PRINTS; PR01084; NAHEXCHNGR.

DR PRINTS; PR01087; NAHEXCHNGR3.

KW Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;

KW Multigene family; Phosphorylation; Polymorphism.

FT DOMAIN 1 11

FT TRANSMEM 12 27

FT DOMAIN 28 59

FT TRANSMEM 60 79

FT DOMAIN 80 81

FT TRANSMEM 82 101

FT DOMAIN 102 110

FT TRANSMEM 111 130

FT DOMAIN 131 134

FT TRANSMEM 135 154

FT DOMAIN 155 180

FT TRANSMEM 181 200

FT DOMAIN 201 209

FT TRANSMEM 210 229

FT DOMAIN 230 249

FT TRANSMEM 250 269

FT DOMAIN 270 298

FT TRANSMEM 299 319

FT DOMAIN 320 339

FT TRANSMEM 340 359

FT DOMAIN 360 366

FT TRANSMEM 367 385

FT DOMAIN 386 435

FT TRANSMEM 436 455

FT DOMAIN 456 832

FT CARBOHYD 325 325

FT VARIANT 144 144

SQ SEQUENCE 832 AA; 92748 MW; 8C8BB7C296CF8740 CRC64;

Query Match 17.2%; Score 475; DB 1; Length 832;

Best Local Similarity 29.1%; Pred. No. 1.le-24;

Matches 148; Conservative 97; Mismatches 196; Indels 68; Gaps 19;

Qy 20 VVALNLFVALLCACIV-LGHLLLEENRMNESITALLGLGTGVTILLISKSSHLVLF- 77

Db 56 IIALWLVASLAKIVFHLSH--KVTSVVPESALLVLGLVGLVLF-----AADHIAST 108

Qy 78 -SEDLFFIYLLPPIIFNAGFOVKKQFPRNVTMLFGAVGTIISCIISLGVTFQFFKL 136

Db 109 LPTPTVFYLLPPIVDAGYFMPNRLFSSNLGSILLVAVGVTVWNAATTGLSLGVFLSG 168

Qy 137 DIGTFDLG--DYLAIGAFAATDSVCTLOVLNQ--DETPLLYSLVFGGVNDATSVVVEN 193

Db 169 IMGEKLTGLDPLFLGSLIAADVPVAVLAVFEEVHVNEVLFIIVFGESLLNDVAVTVLYN 228

Qy 194 AIQSFDTLHNHEAFAHLLGNFLYLLSTLLGAATGLISAYVIKKL-YFGRHSTDEVA 252

Db 229 VFQSVFTLGGDKVTGDCVKGVISFVVVS-LGGTLVGVVFAFLLSLVTRTKHVRVIEP 287

Qy 253 LMMLMAYLSYMLAEFLDLSGILTVFFCGIVGMSHTVHWNTVESSRIITTKHTFATLSFAET 312

Db 288 FVFIISLSYLTSEMLSSLSILAITFCGCCQKYVKANISEQSATTVRYTMKMLASGET 347

Qy 313 FIFYVGMDALIDKWRSDVSTPCTSTAVSISILMGLVMVGRAAFVFPFLSLSLAKKNQS 372

Db 348 IIFMFLGISADVPLIW-----TWTAFVLTLL--FVSVFAIGVVLQTLNLRNMVQL 400

Qy 373 EKINFNMQVVIWMSGLMRGAVSMALAY-----NKFTRAGHTDVRGNAIMTITITVCLFST 428

Db 401 ELID---QVMSYGL-RGAVAFALVALLDGNK-----VREKNLFVSTTIIVVFFTV 448

Qy 429 VVFGMLTKPLISYLLPHONATTSMLSDNTPK-----STHIFLLDQDSFIEPSGNHNPVPRD 485

Db 449 IFQGLTIKPLVQWLKVKR-----SEHREPKLEKLHGRAFD-----HILSAIE 491

QY 486 SIRGLTRPTRTVHY-----WRQFDDSPM 510
 Db 492 DISG-----QIGHNYLKDKNFDRREL 514

RESULT 4
 ID NAH3 RAT STANDARD; PRT; 831 AA.
 AC P26433;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 36, Last annotation update)
 DE SODIUM/HYDROGEN EXCHANGER 3 (NA(+)/H(+) EXCHANGER 3) (NHE-3).
 GN SLC9A3 OR NHE3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
 RX MEDLINE=92250539; PubMed=1577762;
 RA Orlowsky J., Kandasamy R.A., Shull G.E.;
 RT "Molecular cloning of putative members of the Na/H exchanger gene
 family. cDNA cloning, deduced amino acid sequence, and mRNA tissue
 expression of the rat Na/H exchanger NHE-1 and two structurally
 related proteins";
 RT J. Biol. Chem. 267:9331-9339(1992).
 RL
 CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
 CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
 CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
 CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
 CC TRANSDUCTION.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN COLON AND SMALL INTESTINE,
 CC FOLLOWED BY KIDNEY AND STOMACH.
 CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
 CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
 CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
 CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M85300; AAA41702.1; -;
 CC PIR; B40204; B40204.
 CC InterPro; IPR000676; -;
 CC InterPro; IPR001118; -;
 CC Pfam; PF00999; Na_H_Exchange; 1.
 CC PRINTS; PR01084; NAHEXCHNGR.
 CC PRINTS; PR01087; NAHEXCHNGR3.
 CC Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
 CC Multigene family; Phosphorylation.
 KW
 FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 12 25 A (M1) HYDROPHOBIC.
 FT DOMAIN 26 51 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 52 70 B (M2) HYDROPHOBIC.
 FT DOMAIN 71 76 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 77 96 C (M3) (POTENTIAL).
 FT DOMAIN 97 109 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 110 130 D (M4) (POTENTIAL).
 FT DOMAIN 131 136 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 137 157 E (M5) (POTENTIAL).
 FT DOMAIN 158 177 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 178 199 F (M5A) (POTENTIAL).
 FT DOMAIN 200 207 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 208 229 G (M5B) (POTENTIAL).
 FT DOMAIN 230 249 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 250 271 H (M6) (POTENTIAL).
 FT DOMAIN 272 287 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 288 306 I (M7) (POTENTIAL).
 FT DOMAIN 307 337 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 338 359 J (M8) (POTENTIAL).
 FT DOMAIN 360 366 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 367 387 K (M9) (POTENTIAL).
 FT DOMAIN 388 402 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 403 423 L. HYDROPHOBIC.
 FT DOMAIN 424 432 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 433 454 M (M10) (POTENTIAL).
 FT DOMAIN 454 831 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 831 AA; 93105 MW; 77A4BF10DF99E3F CRC64;

Query Match 17.0%; Score 467; DB 1; Length 831;
 Best Local Similarity 28.3%; Pred. No. 3.8e-24;
 Matches 144; Conservative 99; Mismatches 205; Indels 60; Gaps 18;

QY 20 VVALNLFVALLCACIV-LGHLLLENRMNFSITALLIGLGTGVITILLISKGSSHLVLF- 77
 Db 54 IIAIWLIVASLAKIVFHLSH--KVTSPVPESALLIVGLVGLGIV-----WAADHIASFT 106
 QY 78 -SEDLFFIYLLPPIFNAGQVKKQFFRNFTVIMLFGAVGTIISCTIISLGVTFQFKKL 136
 Db 107 LPTLTFEYLLPPIVDAGYFMPNRLFFGNLTGILLYAVIGTIWNAATTGLSLGVFLSG 166
 QY 137 DIGTFDLG--DYLAIGAIFAATDSVCTQLVNLQ-DETPLLYSLVFGEVGVNDATSVVVEN 193
 Db 167 LMGELKIGLLDGLFGLSLAAADVPVAVLAFVEHVHNEVLFIIVFGESLNDAVTVVLN 226
 QY 194 AIGSFOLDTHLNEHAAPFHLGNFLYLFLLSLTLLGAATGLISAYVKKL-YFGRHSTDREVA 252
 Db 227 VFESFVLGDGAVTGVDCVKGIVSFEFVWS-LGCTLVGVIFAPLLSLVTRTKHVRILEPG 285
 QY 253 LMLMAYLSYMLAELELDLSGILVFFCGIVMHSYTHWNTVETSSRTTKHFFATLSLAET 312
 Db 286 FVFVSVLSYLTSEMLSLSLAILAITEFCQCKQYKVNISEQSAATVRYTMKMLASGAET 345
 QY 313 FELYGMDALDIDKVRVSVDTPGTSTAVSSILMGLVMVGVRAAFVPLSFLSLAKKNOS 372
 Db 346 IIFMFLGISAVDPVIM-----TWNTAFVLTLV--FISVTRALGVVLQWILNRYRWQL 398
 QY 373 EKINFNQVVIWMSGLMRGAVSMALAYNKFRAGHTDVRGNAIMITSTITVCLFSTVVF 432
 Db 399 ETID--QVVMYGGGL-RGAVAYALV---VLDEKKVKEKFLFVSTTLIVVFTVIFQ 450
 QY 433 MLTKPLISYLLPQNATTSMLSDNTPK---STHPLLDQDSFIEPSGNHNVPRPDSIRG 489
 Db 451 LTIKPLVQWLKVKR-----SEQRKPKLNEKLUHGRAFD-----HILSAIEDISG 493
 QY 490 FLTRPTRTVHY-----WRQFDDSPM 513
 Db 494 -----QIGHNYLKDKNFDRKFLSKV 515

RESULT 5
 NAH5_HUMAN STANDARD; PRT; 896 AA.
 ID Q14940; Q9Y626;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SODIUM/HYDROGEN EXCHANGER 5 (NA(+)/H(+) EXCHANGER 5) (NHE-5).
 GN SLC9A5 OR NHE5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;

Query Match 16.8%; Score 462; DB 1; Length 834;
 Best Local Similarity 28.6%; Pred. No. 8.4e-24;
 Matches 146; Conservative 97; Mismatches 203; Indels 64; Gaps 18;

Qy 20 VVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTGVTILLISKG---KSSHLV 76
 Db 57 VIALWILVASLAK---IGFHLSHKVSVPESALLVLG-----LVLGGIIVWAADHIA 107

Qy 77 F--SEDFFLYLPPPIIFNAGFQVKKOFFNFVTIMLFGAVGTIISCTIISLGVTFQFK 134
 Db 108 FTLPPTVFFYLLPPIVDAGYFMPNRLFGNLTGILLVAVGVTVWNAATGSLYGVFL 167

Qy 135 KLDIGTDLG--DYLAIGAFAATSDVCTLOVLNQ--DETPLYSLVFGGVVNDATSVV 191
 Db 168 SGLMGDLQIGLDFLFGSLMAAADVPVAVLAVPEEVHNEVLFIIVFGESLNDATVVL 227

Qy 192 FNAQSPDLTHLNEAAFLHGLNLFLLSTLLGAATGLISAYVTKKL-YEGRHSDRE 250
 Db 228 YNVESFVALGGDNTGVCVKGIIVFVVS-LGGFLGVVFAFLLSLVTRTKHVRIE 286

Qy 251 VALMMLMAYLSYMLAEFLDLSGILTFFCGIVMSHYTHNVNTESSRIITKHTFATLSFLA 310
 Db 287 PGFVFIISYLSYLTSEMLSAITATFCGICCKYKANKISEQSATTVRYTMKMASSA 346

Qy 311 ETFFLYVGMADLIDKWRVSVDPTGSTAVSSILMGLVWGEAFAFPPLSLSLNAKN 370
 Db 347 ETILFMFLGISAVNPFIV-----TWNTAFVLLTLV--FISVYRAIGVVLQTLNRYRMV 399

Qy 371 QSEKINFNMQVVIWMSGLMRGAVSMALAYNKFTRAGHTDVRGNAIMITITIVVCLFSTVV 430
 Db 400 QLEPID--QVLSYGGI-RGAAFAVL-----VLLDGDVKVKEKNLWSTIIVVFFTVIF 451

Qy 431 PGMITKPLISYLLPHONATTSMLSDNTPK---SIHIPLLDQDSFTEPGSNHVPSPSI 487
 Db 452 QGLTIKPLVOWLKVKR-----SEHREPLNEKLHGRAPD-----HILSAIEDI 494

Qy 488 RGLTRTRTVHYV---WRQFDDSFMRPV 513
 Db 495 SG-----QIGHNLYRDKWSHEDRKFSLRV 518

RESULT 8
 NAH3_DIDMA STANDARD; PRT; 839 AA.
 AC Q28362;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SODIUM/HYDROGEN EXCHANGER 3 (NA(+)/H(+) EXCHANGER 3) (NHE-3).
 GN SLC9A3 OR NHE3.
 OS Didelphis marsupialis virginiana (North American opossum).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
 OC NCBI_TaxID=9267;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95358256; PubMed=7631739;
 RA Amemiya M., Yamaji Y., Cano A., Moe O.W., Alpern R.J.;
 RT "Acid incubation increases NHE-3 mRNA abundance in OKP cells.";
 RL Am. J. Physiol. 269:C126-C133(1995).
 CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
 CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
 CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
 CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
 CC TRANSDUCTION.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
 CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
 CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
 CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L42522; AAA98816.1; -;
 CC InterPro; IPR001118; -;
 CC Pfam; PF00999; Na.H.Exchanger; 1.
 CC PRINTS; PR01084; NAHEXCHNGR.
 CC PRINTS; PR01087; NAHEXCHNGR3.
 CC Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
 CC Multigene family; Phosphorylation.
 CC FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 14 31 M1 (POTENTIAL).
 CC FT DOMAIN 32 65 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 66 85 M2 (POTENTIAL).
 CC FT DOMAIN 86 87 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 88 107 M3 (POTENTIAL).
 CC FT DOMAIN 108 113 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 114 133 M4 (POTENTIAL).
 CC FT DOMAIN 134 150 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 151 170 M5 (POTENTIAL).
 CC FT DOMAIN 171 186 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 187 206 M5A (POTENTIAL).
 CC FT DOMAIN 207 215 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 216 235 M5B (POTENTIAL).
 CC FT DOMAIN 236 255 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 256 276 M6 (POTENTIAL).
 CC FT DOMAIN 277 299 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 300 319 M7 (POTENTIAL).
 CC FT DOMAIN 320 347 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 348 366 M8 (POTENTIAL).
 CC FT DOMAIN 367 370 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 373 392 M9 (POTENTIAL).
 CC FT DOMAIN 393 442 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 443 462 M10 (POTENTIAL).
 CC FT DOMAIN 463 839 CYTOPLASMIC (POTENTIAL).
 CC FT CARBOHYD 331 331 N-LINKED (GLCNAC...) (POTENTIAL).
 CC SEQUENCE 839 AA; 94765 MW; F6P9EF778D5DBB2 CRC64;

Query Match 16.7%; Score 459; DB 1; Length 839;
 Best Local Similarity 30.9%; Pred. No. 1.3e-23;
 Matches 136; Conservative 86; Mismatches 168; Indels 50; Gaps 16;

Qy 20 VVALNLFVALLCACIVLGHLEENRWNNESIT-----ALLIGLGTGVTILLISKG---K 70
 Db 62 IIALWILVASLAK--IVFHL-----SHKVSVPESALLVIG-----LILGIWA 106

Qy 71 SSSLVFP--SEDLFFIYLLPPIIFNAGFQVKKOFFNFVTIMLFGAVGTIISCTIISLG 128
 Db 107 ADHIASFLLTPTVFFYLLPPIVDAGYFMPNRLFGNLTGILLVAVIGTVWNAATFGLS 166

Qy 129 VTQFFKKLDIGTDLG--DYLAIGAFAATSDVCTLOVLNQ--DETPLYSLVFGGVVND 185
 Db 167 LYGYLVSIGMGDLISGLLDFLFGSLAAADVPVAVLAVPEEVHNDVLFIIVFGESLND 226

Qy 186 ATSVVFNALQSPDLTHLNEAAFLHGLNLFLLSTLLGAATGLISAYVTKKL-YFGR 244
 Db 227 AVTVVLNVDSFVSFGADKVTGVCVKGIIVFVVS-LGGTLIGIIFAFLLSLVTRFTK 285

Qy 245 HSTOREVALMMLMAYLSYMLAEFLDLSGILTFFCGIVMSHYTHNVNTESSRIITKHTFA 304
 Db 286 HVRIEFGFVFIISYLSYLTSEMLSAITATFCGICCKYKANKISEQSATTVRYTMK 345

Qy 305 TLSLFAETFFLYVGMADLIDKWRVSVDPTGSTAVSSILMGLVWGEAFAFPPLSL 362
 Db 346 MLASGAETIIFMFLGISAVNPFIV-----TWNTAF-----ILLTLVFSYVRAIGVVLQTW 396

Qy 363 LSNLAKNQSEKINFNMQVVIWMSGLMRGAVSMALAYNKFTRAGHTDVRGNAIMITIT 422

QY 352 GRAAFVPLSLAKNQSEKINFMQVVIWWSGLMRGAVSMALAY-----NKETRAGH 407
Db 405 WRALGVFLTVQVNWFR---TIPLTFKQDFIITAYGGL-RGAICFALVFLPATVFPFR--- 457
QY 408 TDVRGNAMITSTITVCLSTVVFGLMTPLISYL-LPHQNAATSMLSDDNTPKSHIPL 466
Db 458 -----KKLFITAAVIVFFITGLTIRPLVEFLDVKRSNKQQAQVSEE-----IHCRRF 507
QY 467 LDQ-DSFIEPSGNHNVPRDSIRGFLTRPTVHYWYR-----QFDDSEMR 511
Db 508 FDHVKTGIEDVCGH-----WGHNFWRDKFKFDKDKYL 540
RESULT 10
NAH2_RABIT STANDARD; PRT; 809 AA.
AC P50482;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 2 (NA(+)/H(+)) EXCHANGER 2) (NHE-2).
GN SLC9A2 OR NHE2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEW ZEALAND WHITE;
RX MEDLINE=93280159; PubMed=7685025;
RA Tse C.-M., Levine S.A., Yun C.H., Montrose M.H., Little P.J.,
RA Pouyssegur J., Donowitz M.;
RT "Cloning and expression of a rabbit cDNA encoding a serum-activated
RT ethylisopropylamide-resistant epithelial Na+/H+ exchanger isoform
RT (NHE-2).";
RL J. Biol. Chem. 268:11917-11924(1993).
CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN
CC COLONIC SODIUM ABSORPTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: HIGH LEVELS IN INTESTINE AND KIDNEY.
CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L13733; ; NOT_ANNOTATED_CDS.
DR InterPro; IPR000676; .
DR Pfam; PF00999; Na.H.Exchanger; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01086; NAHEXCHNGR2.
KW Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW Multigene family; Phosphorylation.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 37 A (M1) (POTENTIAL).
FT DOMAIN 38 79 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 80 100 B (M2) (POTENTIAL).
FT DOMAIN 101 106 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 107 127 C (M3) (POTENTIAL).
FT DOMAIN 128 138 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 139 159 D (M4) (POTENTIAL).
FT DOMAIN 160 168 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 169 189 E (M5) (POTENTIAL).
FT DOMAIN 190 208 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 209 229 F (M5A) (POTENTIAL).
FT DOMAIN 230 236 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 237 257 G (M5B) (POTENTIAL).
FT DOMAIN 258 277 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 278 298 H (M6) (POTENTIAL).
FT DOMAIN 299 307 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 308 328 I (M7) (POTENTIAL).
FT DOMAIN 329 360 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 361 381 J (M8) (POTENTIAL).
FT DOMAIN 382 391 K (M9) (POTENTIAL).
FT TRANSMEM 392 412 L (M10) (POTENTIAL).
FT DOMAIN 413 429 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 430 450 L (M10) (POTENTIAL).
FT DOMAIN 451 458 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 459 479 M (M11) (POTENTIAL).
FT DOMAIN 480 809 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 809 AA; 90744 MW; DBD00B45443D87A6 CRC64;
Query Match 16.5%; Score 454.5; DB 1; Length 809;
Best Local Similarity 28.4%; Pred. No. 2.6e-23;
Matches 150; Conservative 94; Mismatches 206; Indels 79; Gaps 22;
QY 7 SKLPSLSTSDHASV---VALNLFVALLACIVLGHLEE-NRWNNESITALLIGLGTGT 62
Db 66 SRLP-VFTLDYPHQVQIPFEITLWILLASLAKIGFLYHKLPTIVPESCLIMVGLLGGI 124
QY 63 ILLSKGSSHLLVSEDFLYLLPPIFNFAQVKKKOFFRNFTVIMLFGAVGTIISC 122
Db 125 IFGVDE-KSPAM--KTDVFLYLLPPIVDAGYFMFTRFFENLGHIFVAVVGTUWS 181
QY 123 TIIISLGVTFKKLDIGTDLGDY-----LAIGAIFAATSDVCTQLVL-NODETPLYSL 176
Db 182 --IGIGVS-LFGICQIEAFGLSDITLLQNLFLGSLISAVDPVAVLAVFENHVNQYL 238
QY 177 VFGGVNDATSVVFNAIQSF-DLTHLNEAFHLLGNFLYLLFLLSTILGAATGLISAY 235
Db 239 VFGESLNDATVVVLYNLFSQCMKTETIDVFAGIANFVVGIGVGLIGLGFIAAF 298
QY 236 VIKKLYGRHSTDREVALMMLMAYLSYMLAEFLDLSGLITVFFGIVMSHYTHVNWTESS 295
Db 299 TTR---ETHNIRVIEPLFVLYSLVSYTAEMFHLSIMAITACAMTNKYVEENSQKS 355
QY 296 RITTKHFTATLSLAETFIIFYVGMALDID---KWRVSVDTPGTSTAVSSILMGLVMVG 352
Db 356 YTTIKYFMKMLSSVSETLIFMGVSTVGKNHEWNAFVCF-----LAFCLIW 404
QY 353 RAAVPEPLSLNLAKKNQSEKINFMQVVIWWSGLMRGAVSMALAY-----NKETRAGH 408
Db 405 RALGVFLTVQVNWFR---TIPLTFKQDFIITAYGGL-RGAICFALVFLPATVFPFR--- 456
QY 409 DVRGNAIMSTITVCLSTVVFGLMTPLISYL-LPHQNAATSMLSDDNTPKSHIPL 467
Db 457 ----KKLFITAAVIVFFITGLTIRPLVEFLDVKRSNKQQAQVSEE-----IHCRRF 507
QY 468 DQ-DSFIEPSGNHNVPRDSIRGFLTRPTVHYWYR-----QFDDSEMR 511
Db 508 DHVKTGIEDVCGH-----WGHNFWRDKFKFDKDKYL 539
RESULT 11
NAH2_HUMAN STANDARD; PRT; 812 AA.
ID NAH2_HUMAN AC Q90BY0;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 2 (NA(+)/H(+)) EXCHANGER 2) (NHE-2).
GN SLC9A2 OR NHE2.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=99375108; PubMed=10444453;
RA Malakooti J., Dahdal R.Y., Schmidt L., Layden T.J., Dudeja P.K.,
RA Ramaswamy K.;
RT "Molecular cloning, tissue distribution, and functional expression of
RT the human Na(+)/H(+) exchanger NHE2";
RL Am. J. Physiol. 277:G383-G390(1999).
RN (2)
RP SEQUENCE FROM N.A.
RA Hou S., Wohldmann P.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC SODIUM ION CHEMICAL GRADIENT. SEEMS TO PLAY AN IMPORTANT ROLE IN
CC COLONIC SODIUM ABSORPTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL MUSCLE, COLON AND
CC KIDNEY. LOWER LEVELS IN THE TESTIS, PROSTATE, OVARY, AND SMALL
CC INTESTINE.
CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC -!- SIMILARITY: BELONGS TO THE Na(+)/H(+) EXCHANGER FAMILY.
CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC DOMAINS IN THE Na(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
ER EMBL; AF073299; AAD41635.1; -;
DR EMBL; AC007239; AAF19248.1; -;
DR MIM; 600530; -;
DR InterPro; IPR000676; -;
DR InterPro; IPR001953; -;
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01086; NAHEXCHNGR2.
KW Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW Multigene family; Phosphorylation.
FT DOMAIN 1 13
FT DOMAIN 14 33
FT DOMAIN 34 79
FT DOMAIN 80 100
FT DOMAIN 101 106
FT DOMAIN 107 127
FT TRANSMEM 128 138
FT DOMAIN 139 159
FT DOMAIN 160 168
FT TRANSMEM 169 189
FT DOMAIN 190 208
FT TRANSMEM 209 229
FT DOMAIN 230 236
FT TRANSMEM 237 257
FT DOMAIN 258 277
FT TRANSMEM 278 298
FT DOMAIN 299 307
FT TRANSMEM 308 328
FT DOMAIN 329 360
FT TRANSMEM 361 381
FT DOMAIN 382 391
FT TRANSMEM 392 412
FT DOMAIN 413 429
FT DOMAIN 430 450
FT DOMAIN 451 458
EXTRACELLULAR (POTENTIAL).
A (M1) HYDROPHOBIC.
CYTOPLASMIC (POTENTIAL).
B (M2) HYDROPHOBIC.
CYTOPLASMIC (POTENTIAL).
C (M3) (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
D (M4) (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
E (M5) (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
F (M5A) (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
G (M5B) (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
H (M6) (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
I (M7) (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
J (M8) (POTENTIAL).
CYTOPLASMIC (POTENTIAL).
K (M9) (POTENTIAL).
EXTRACELLULAR (POTENTIAL).
L HYDROPHOBIC.
EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 459 479 M13 (POTENTIAL).
FT DOMAIN 480 812 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 812 AA; 91519 MW; 17BE177DC3830D0A CRC64;

Query Match 16.4%; Score 452.5; DB 1; Length 812;
Best Local Similarity 28.4%; Pred. No. 3.5e-23;
Matches 150; Conservative 96; Mismatches 204; Indels 79; Gaps 22;

QY 7 SKLPSTSDHASV---VALNLFVALLCACIVLGHLLLE- NRWMNESITALLIGLGTGT 62
DB 66 SRLP-VFTLDYPHVQPPFEITLLWLLASLAKGFHYHLKPTIVPESCLLIWGLLGI 124
QY 63 ILLISKSKSHLLVFSSEDLFFIYLLPPIIFNAGFOVKKKQFFRNFVTIMLFGAVGTIISC 122
DB 125 IFGVDE-KSPAPAM-KTDVEFLYLLPPIVDAGYFMPTRPFENIGTIPWYAVVGTLWNS 181
QY 123 TITSLGVTOFFKKLDIGTFDLGDY-----LATGAIPAATDSVCTLOVL-NQDETPLLISL 176
DB 182 --IGIGVS-LFGICQTEAFGLSDITLLQNLFFGLISAVDPVAVLAVFENIHVNEQLYIL 238
QY 177 VEGEYVNDATSVVFNATQSF-DLTHLNHEAAFHLLGNFLYFLLLSTLLGAATGLISAY 235
DB 239 VFGESLNDATVVLNLFKSCQMTIETIDVFAGIANFFVVGIGVLIGIFLGFIAAF 298
QY 236 VIRKLYFGRHSTDREVALMMLMAYLSMLAEFLDLSGILTFFCGIVMGSHTYTHNVTESS 295
DB 299 TTR--FTHNIRVIEPLFVFLYSYLSYITAEMHLSEIMAITACAMTMNKYVEENVSKS 355
QY 296 RITTKITFATLSPLAETFTFLYVGMALDID---KWRVSVDTPGTSIAVSSILMLGVWVG 352
DB 356 YTTIKYFMKMLSSVSETLFIEMGVSTGVKNHEWNMAFVCF-----LAFCLMW 404
QY 353 RAAFPVPLSLNLAKKQSEKINENQVIMVWVSGILMRGAVSMALAY----NKFTRAGHT 408
DB 405 RALGVFLVQVIN---RFRIPITFKDQIIAYGGL-RGAICFALVFLPAAVFP----- 456
QY 409 DVRGNAMITSTITVCLFSTVFEGLTKPLISYL-LPHQNATTSMLSDNTPKSIHPLL 467
DB 457 ----KKLITAAIWIFFTVFILGITIRPLVEFLDKRNSKKQAVSEE-----IYCRLF 507
QY 468 DQ-DSIEPESGNHVPDPDSIRGFLRPTVHYVWR---QFDDSFMR 511
DB 508 DHVKTGIEDVCGH-----WGHNFWDRKFKFPDDKYL 539

RESULT 12
NAHL_HUMAN STANDARD; PRT; 815 AA.
AC P19634;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+) EXCHANGER 1) (NHE-1) (NA+/H+
DE ANTIporter, AMILORIDE-SENSITIVE) (APNH).
GN SLC9A1 OR NHE1 OR APNH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=89106219; PubMed=2536298;
RA Sargent C., Franchi A., Pouyssegur J.;
RT "Molecular cloning, primary structure, and expression of the human
RL growth factor-activatable Na+/H+ antiporter.";
RN Cell 56:271-280(1989).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=90140739; PubMed=2154036;

RA Sardet C., Counillon L., Franchi A., Pouyssegur J.;
RT "Growth factors induce phosphorylation of the Na⁺/H⁺ antiporter,
RL glycoprotein of 110 kD";
RN Science 247:723-726(1990).
RP [3]
RA REVISIONS.
RX MEDLINE-91293066; PubMed-1712287;
RA Tse C.-M., Ma A.I., Yang W.W., Watson A.J.M., Levine S.,
RA Montrose M.H., Potter J., Sardet C., Pouyssegur J., Donowitz M.;
RT "Molecular cloning and expression of a cDNA encoding the rabbit ileal
RT villus cell basolateral membrane Na⁺/H⁺ exchanger";
RL EMBO J. 10:1957-1967(1991).
RN [4]
RA SEQUENCE FROM N.A.
RP TISSUE=Heart;
RC MEDLINE-94111706; PubMed-8283968;
RX Fliegel L., Dyck J.R., Wang H., Fong C., Haworth R.S.;
RA "Cloning and analysis of the human myocardial Na⁺/H⁺ exchanger";
RT Mol. Cell. Biochem. 125:137-143(1993).
RN [5]
RA SEQUENCE FROM N.A.
RP MEDLINE-20375279; PubMed-10913675;
RX Garden O.A., Musk P., Worthington-White D.A., Dewey M.J., Rich I.N.;
RA "Silent polymorphisms within the coding region of human
RT sodium/hydrogen exchanger isoform-1 cDNA in peripheral blood
RT mononuclear cells of leukemia patients: A comparison with healthy
RT controls";
RL Cancer Genet. Cytogenet. 120:37-43(2000).
CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC TRANSDUCTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: KIDNEY AND INTESTINE.
CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC -!- MISCELLANEOUS: INHIBITED BY AMILORIDE AND 5-AMINO-SUBSTITUTED
CC DERIVATIVES AND ACTIVATED IN A COOPERATIVE FASHION BY
CC INTRACELLULAR H⁺. FULLY ACTIVE AT ACIDIC PH. THE ANTIPORTER IS
CC VIRTUALLY TURNED OFF AT NEUTRAL PH. IN QUIESCENT CELLS UPON GROWTH
CC FACTOR STIMULATION, THE APPARENT AFFINITY FOR INTERNAL H⁺ IS
CC INCREASED, RESULTING IN A PERSISTENT RISE IN CYTOPLASMIC PH.
CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+) EXCHANGER FAMILY.
CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC DOMAINS IN THE NA(+)/H(+) EXCHANGERS VARY AMONG AUTHORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81768; AAF59460.1; ALT_SEQ.
DR EMBL; S68616; AAC60606.1; -
DR EMBL; AAF141350; AAF21350.1; -
DR EMBL; AAF141351; AAF21351.1; -
DR EMBL; AAF141352; AAF21352.1; -
DR EMBL; AAF141353; AAF21353.1; -
DR EMBL; AAF141354; AAF21354.1; -
DR EMBL; AAF141355; AAF21355.1; -
DR EMBL; AAF141356; AAF21356.1; -
DR EMBL; AAF141357; AAF21357.1; -
DR EMBL; AAF141358; AAF21358.1; -
DR EMBL; AAF141359; AAF21359.1; -
DR EMBL; AAF146430; AAF25592.1; -
DR EMBL; AAF146431; AAF25593.1; -
DR EMBL; AAF146432; AAF25594.1; -
DR EMBL; AAF146433; AAF25595.1; -
DR EMBL; AAF146434; AAF25596.1; -
DR EMBL; AAF146435; AAF25597.1; -
DR EMBL; AAF146436; AAF25598.1; -

DR EMBL; AFI146437; AAF25599.1; -
DR EMBL; AFI146438; AAF25600.1; -
DR EMBL; AFI146439; AAF25601.1; -
DR PIR; A31311; A31311.
DR MIM; 107310; -
DR InterPro; IPR000676; -
DR InterPro; IPR001970; -
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01085; NAHEXCHNGR1.
KW Transmembrane; Glycoprotein; Sodium transport; Symport;
KW Multigene family; Phosphorylation.
FT DOMAIN 1 15
FT TRANSMEM 16 35
FT DOMAIN 36 107
FT TRANSMEM 108 127
FT DOMAIN 128 129
FT TRANSMEM 130 149
FT DOMAIN 150 154
FT TRANSMEM 155 174
FT DOMAIN 175 191
FT TRANSMEM 192 211
FT DOMAIN 212 227
FT TRANSMEM 228 247
FT DOMAIN 248 256
FT TRANSMEM 257 276
FT DOMAIN 277 294
FT TRANSMEM 295 315
FT DOMAIN 316 338
FT TRANSMEM 339 358
FT DOMAIN 359 386
FT TRANSMEM 387 406
FT DOMAIN 407 410
FT TRANSMEM 411 430
FT DOMAIN 431 480
FT TRANSMEM 481 500
FT DOMAIN 501 815
FT CARBOHYD 75 75
FT CARBOHYD 370 370
SQ SEQUENCE 815 AA; 90763 MW; 02EC748C79DF6526 CRC64;

Query Match 16.1%; Score 443; DB 1; Length 815;
Best Local Similarity 26.0%; Pred. No. 1.5e-22;
Matches 151; Conservative 99; Mismatches 205; Indels 126; Gaps 22;

Qy 23 LNFVALICACIVLGHLEE-NRWNESITALLIGLGTGVITLLISKGKSHLLVFSDEL 81
Db 105 ISLWLLACLAKMGHVIPTISSIVPESCLLVVGLVGGI-----KGVETPPFLQSDV 160
Qy 82 FFYILLPIIPNAGFOVKKQFFNFVTIMLFGAVGI-----ISCTIISLGVTOFF 133
Db 161 FFLFLLPIILDAGYFLPLROFTENLGTILFAVVGTLWNAFFLGLGLMYAVCLVGGEQ-- 218
Qy 134 KKLDIGTFDLGDLAIGAIFAATSDVCTLOVLNODE-TPLLYSLVFCGEGVVDATSVVVF 192
Db 219 ----INNIGLLDNLFGSISAVDPVAVLAFVEIHINELLHLVFGESLNDKAVTVVLY 274
Qy 193 NAIQSFDLTHLHBAFAFHLNGLNLYLFLLS---TLLGAATGLISAYVIKKLYFGHSTDR 249
Db 275 HLFEF--ANYEHVGIVDIFLGLFSFFVGLVGVGVIAAFTSR---FTSHIRVI 329
Qy 250 EVALMMLMAYLSYMLAEFLDLSGLITVFFGIVGVMSHTVWNVTESSRITTKHFTATLSFL 309
Db 330 EPLFVFLYSYMAVLSAELFHLSGIMALIASGVVMRPVEANISHKSHTTIKYFLKMWSSV 389
Qy 310 AETIFLYVGMALDIDKWSVSDPTCT-----SIAVSIILMGLVMVGRAAFVPLSFLS 364
Db 390 SETLIFILG-----VSTVAGSHWNWTFVISTLL--FCLIRVILVGLGLTWFI 436
Qy 365 NLAQKNQSEKINFNMQVYIWMVSGLMRGAVSMALAY-----NKFTRAGHTDVRGNAMITST 420
Db 437 N---KFRIVKLTPKDQFIAYGGL-RCAIAFSLGLYLLDKKHPMCO-----LFLTAI 484


```

Db 437 N---KFRIVLTKPKDQFIAYGGL- RGAIAFSLGYLLDKKHPMCD-----LFLTAI 484
QY 421 ITVCLFSTVVGMLTKPLSYLLPHONATTSMLSDNTPKSHIPLDDODSFIEPSGNHN 480
Db 485 ITVIFTFVVGQMTIRPLVDLLAVKKQETKR-----SINEEIHQFLD-----HL 530
QY 481 VPRPDSIRGFLTRPTRTVHYWYR-----QFDDSFMRPVF----- 514
Db 531 LTGIEDICGHYG-----HHWKDKLNRRNKYVKKCLLAGERSKEPQLIAFYHKMEMQ 584
QY 515 -----GGRGFVPFVPGS-----PTERNPPDLSK 537
Db 585 AIELVESGGMGKIPSAVSTVSMQNIHPKALPAERILPALSK 625

RESULT 14
NAHL_PIG
ID NAHL_PIG STANDARD; PRT; 818 AA.
AC P48762;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SODIUM/HYDROGEN EXCHANGER 1 (NA(+)/H(+)) EXCHANGER 1) (NHE-1).
GN SLC9A1 OR NHE1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=92087905; PubMed=1661081;
RX Reilly R.F., Hildebrandt F., Blemesderfer D., Sardet C.,
RA Pouyssegur J., Atkinson P.S., Blayman C.W., Igarashi P.;
RT "cDNA cloning and immunolocalization of a Na(+)-H+ exchanger in
RT LLC-PK1 renal epithelial cells.";
RL Am. J. Physiol. 261:F1088-F1094(1991).
CC -!- FUNCTION: INVOLVED IN PH REGULATION TO ELIMINATE ACIDS GENERATED
CC BY ACTIVE METABOLISM OR TO COUNTER ADVERSE ENVIRONMENTAL
CC CONDITIONS. MAJOR PROTON EXTRUDING SYSTEM DRIVEN BY THE INWARD
CC SODIUM ION CHEMICAL GRADIENT. PLAYS AN IMPORTANT ROLE IN SIGNAL
CC TRANSDUCTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- PTM: PHOSPHORYLATED (POSSIBLE).
CC -!- SIMILARITY: BELONGS TO THE NA(+)/H(+)) EXCHANGER FAMILY.
CC -!- CAUTION: THE NUMBER, LOCALIZATION AND DENOMINATION OF HYDROPHOBIC
CC DOMAINS IN THE NA(+)/H(+)) EXCHANGERS VARY AMONG AUTHORS.
CC -!- CAUTION: HYDROPHOBIC DOMAINS A, B AND L ARE NOT BELIEVED TO BE
CC TRANSMEMBRANAL, BUT ONLY MEMBRANE-ASSOCIATED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M89631; AAA31092.1; -.
DR EMBL; S71135; AAB20633.1; -.
DR InterPro; IPR0000576; -.
DR InterPro; IPR001970; -.
DR Pfam; PF00999; Na_H-Exchanger; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01085; NAHEXCHNGR.
KW Transmembrane; Glycoprotein; Sodium transport; Transport; Symport;
KW Multigene family; Phosphorylation.
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 12 31 A (M1) HYDROPHOBIC.
FT DOMAIN 32 101 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 102 123 B (M2) HYDROPHOBIC.
FT DOMAIN 124 126 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 127 146 C (M3) (POTENTIAL).

```

```

FT DOMAIN 147 158 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 159 179 D (M4) (POTENTIAL).
FT DOMAIN 180 184 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 185 206 E (M5) (POTENTIAL).
FT DOMAIN 207 226 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 227 247 F (M5A) (POTENTIAL).
FT DOMAIN 248 256 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 257 278 G (M5B) (POTENTIAL).
FT DOMAIN 279 297 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 298 318 H (M6) (POTENTIAL).
FT DOMAIN 319 332 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 333 353 I (M7) (POTENTIAL).
FT DOMAIN 354 384 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 385 406 J (M8) (POTENTIAL).
FT DOMAIN 407 412 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 413 434 K (M9) (POTENTIAL).
FT DOMAIN 435 448 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 449 469 L, HYDROPHOBIC.
FT DOMAIN 470 478 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 479 499 M (M10) (POTENTIAL).
FT DOMAIN 500 818 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 370 370 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 683 683 H -> Y (IN AAB20633).
SQ SEQUENCE 818 AA; 90987 MW; 9329F7D9A51D3DC9 CRC64;

Query Match 15.8%; Score 436; DB 1; Length 818;
Best Local Similarity 25.5%; Pred. No. 4.5e-22;
Matches 148; Conservative 105; Mismatches 202; Indels 126; Gaps 22;

QY 23 LNLFWALLCACIVLGHLLLE-NRWNNESITALLGLGTGVTILLISKSKSHLLVFSEDL 81
Db 105 ISLWLLACLKMGIFHVIPTISSIPESCLLVGLVGLIKAV--GETPPFL--OSEV 160
QY 82 FFYLLPPIIPNAGFOVKKQFFNFVTIMLFAGVGTI-----ISCTIISLGVTQFF 133
Db 161 FFELLLPIILDAGYFLPLRQFTENLGTILFAVVGTLWNAFFLGLMYAVCLVGGEQ-- 218
QY 134 KKLDTGIFDLGDYLAIGAFAAFTASVCTQLVNODE--TPLLYSLVFGEGVVDNATSVVF 192
Db 219 ----INNIGLLDNLFGSIISAVDPVAVLFAVEEIHLLHLLVFGESLINDAVTVLY 274
QY 193 NAIQSFDLTHLNEAFAFLHCLNGLYFLFLS---TLGAAATGLISAYVIKKLYGHRHSTD 249
Db 275 HLFEF--ANYDRVIGVDIVLGLFSLFVSVLGGVGVGVIAAFTSR---FTSHIRVI 329
QY 250 EVALMLMAYLSYMLABELDLISGLITVFFCGIVMNSHYTHWNTESSRITTKHTFATLSFL 309
Db 330 EPLEFVLYSYMAYLSABELHLISGLIMALIASGVVMRPYVEANISHKSHTTIKYFLKMWSV 389
QY 310 AETFIFLVGMADALDIDKRSVSDTPCT----SIAVSSILMGLVMYGRAAFVPLSFLS 364
Db 390 SETLIFILFG-----VSTVAGSHHWNWTFVISTILL--FCLIARVLGLGLTWFI 436
QY 365 NLAKKNQSEKINENMQVIVWMSGLMRGAVSMALAY----NKFTRAGHTDVRGNAIMTST 420
Db 437 N---KFRIVLTKPKDQFIAYGGL- RGAIAFSLGYLLDKKHPMCD-----LFLTAI 484
QY 421 ITVCLFSTVVGMLTKPLSYLLPHONATTSMLSDNTPKSHIPLDDODSFIEPSGNHN 480
Db 485 ITVIFTFVVGQMTIRPLVDLLAVKKQETKR-----SINEEIHQFLD-----HL 530
QY 481 VPRPDSIRGFLTRPTRTVHYWYR-----QFDDSFMRPVF----- 514
Db 531 LTGIEDICGHYG-----HHWKDKLNRRNKYVKKCLLAGERSKEPQLIAFYHKMEMQ 584
QY 515 -----GGRGFVPFVPGS-----PTERNPPDLSK 537
Db 585 AIELVESGGMGKIPSAVSTVSMQNIHPKALPAERILPALSK 625

RESULT 15
NAHL_ONCMY

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: November 18, 2001, 08:47:25 ; Search time 2907.92 Seconds
(without alignments)
11585.173 Million cell updates/sec

Title: US-09-271-584A-1
Perfect score: 2178
Sequence: 1 cctctctgttctgtctctg.....aaaaaaaaaaaaaaaaaaaa 2178

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_p11.*
13: gb_p12.*
14: gb_p13.*
15: gb_p14.*
16: em_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pi.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vi.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_v12.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_ro1.*
95: gb_ro2.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB	ID	Description
1	1615.8	74.2	1619	12	AF056190	Arabidops
2	1614	74.1	1614	12	AF106324	Arabidops
3	832.8	38.2	2637	12	AB038492	Atriplex
4	774.8	35.6	2218	12	AB033989	Ipomoea n
5	768.8	35.3	2313	12	AB021878	Oryza sat
6	693	31.8	90019	15	TM021B04	Arabidops
7	266.6	12.2	600	14	AY028416	Citrus x
8	229.6	10.5	421	13	AF279671	Mesembrya

Db 721 TATACTTTGGAGGCACTCAACTGACGGAGAGGTTGCCCTTATGATGCTTATGGCGTATC 780
Qy 1064 ttcttatctgtcgtgagccttttcgaacttgagcgtatccctcaactgtgtttttctgtg 1123
Db 781 TTTCTTATATGCTTGTGCTTTCACCTTGAGCGGTATCCACACTGTCTTTTCTGTG 840
Qy 1124 gtattgtatgtccattatcacatggccacaaatgtaacggagagctcaagaataacaacaa 1183
Db 841 GTATTGTGATGCCATTACATATGGCACAAATGTAACGGAGAGCTCAAGAAATACACAA 900
Qy 1184 agcatacctttgcaactttgttcatttttcgttcggagacatttattttgtatgttggaa 1243
Db 901 AGCATACCTTTGGCAACTTTTGTCACTTCTCGGAGAGCAATTTATTTCTGTATGTGGAA 960
Qy 1244 tggatgcttggacattgacaagtggagatccgtgagtgacacacggagaacatcatcg 1303
Db 961 TGGATGCTTGGACATTTGACAAGTGGAGATCCGTGAGTGACACCCGGGAACATCATCG 1020
Qy 1304 cagtgaactcaactcaactaatgggtgtggttcgtatggttgaagagcagcgttcgtttccgt 1363
Db 1021 CAGTGAAGCTCAATCTTAATGGGTGTGTCATGCTTGGAGAGCAGCGTTCGTCTTCCGT 1080
Qy 1364 tategtttctataacttagccaagaagaatacaaaagcgagaaatacaactttaacatgc 1423
Db 1081 TATCGTTTCTATCTAACTTAGCCCAAGAAGAAATCAAGCCGAGAAATCAACTTTAACA 1140
Qy 1424 aggttggatttgggtgctggtctcatgagaggtgctgtatctatgctgtccttgcataca 1483
Db 1141 AGGTTGATGTTGGTGTGCTGGTCTCATGAGAGTGCTGTATCTATGGCTCTTGCATACA 1200
Qy 1484 acaagtttaacaaagggccgggacacacagatgtagcggggaatgcaatcatgatcaagtag 1543
Db 1201 ACAAGTTTACAAGGGCCGGSCACACAGATGTACCGGGAATGCAATCATGATCACAGTA 1260
Qy 1544 cgataactgtctgtcttttagcacagtgtgttgtagtgatgctgacaaacacactataa 1603
Db 1261 CGATACTGTCTCTTTTGTAGCACAGTGTGTTGGTATGCTGACCAACCACTCATATA 1320
Qy 1604 gctactattaccgaccagaaagccaccagcagcagtggttctatctgatgacaaacaccccaa 1663
Db 1321 GCTACCTATTACCGCACCAAGAGCCACCACGAGCATGTTATCTGATGACAAACCCCAA 1380
Qy 1664 aatccatacatcctttgttggaccacagactcgttcattagccttcagggaaccca 1723
Db 1381 AATCCATACATATCCCTTTTGGACCAAGACTCGCTTCATTGAGCCTTCAGGGGAACCA 1440
Qy 1724 atgtgctcggcgtgacagtatacgtggtctcttgacacggccactcgaaacgctgcatt 1783
Db 1441 ATGTGCTCGGCTGACATGATACGTGGCTTCTTGACACGGCCCACTCGAACCGTGCA 1500
Qy 1784 actactggagacaatttgatgactccttcacgtgaccccgcttttgagggtcgtggtttg 1843
Db 1501 ACTACTGGAGACAATTTGATGACTCTTTTCATGCGACCCGCTTTTGAGGTCGTGGCTTG 1560
Qy 1844 taccctttgtcaggttcccaactgagagaacccctcctgatatttagaaggttga 1902
Db 1561 TACCCTTTGTTCAGGTTCTCCAAGTGTAGAGAAACCTCCTGATCTTAGTAAGGCTTGA 1619

RESULT 2
AF106324
LOCUS
DEFINITION
Arabidopsis thaliana sodium proton exchanger Nhxl mRNA, partial cds.
ACCESSION
AF106324
VERSION
AF106324.1 GI:4324596
KEYWORDS
SOURCE
thale cress.
ORGANISM
Arabidopsis.thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 1614)

AUTHORS Gaxiola,R.A., Rao,R., Sherman,A., Grisafi,P., Alper,S.L. and Fink,G.R.
TITLE The Arabidopsis thaliana proton transporters, AtNhxl and Avp1, can function in cation detoxification in yeast
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (4), 1480-1485 (1999)
MEDLINE 99145575
REFERENCE 2 (bases 1 to 1614)
AUTHORS Gaxiola,R.A., Rao,R., Sherman,A., Grisafi,P., Alper,S.L. and Fink,G.R.
TITLE Direct Submission
JOURNAL Submitted (12-NOV-1998) Whitehead, Nine Cambridge Center, Cambridge, MA 02142, USA
FEATURES
Location/Qualifiers
source 1..1614
CDS
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
1..>1614
/codon_start=1
/product="sodium proton exchanger Nhxl"
/protein_id="AAD16946.1"
/translation="MLDSLVSKLPSLSTSDHASVVALNLFVALICACIVLHLEENRWMNESITALLIGLGTVTIILLISKGKSHLLVFSEDLFFIYLLPPIIFNAGFQVKKQFFRNFVITMLFGAVGTIISCTIISLGVTFQFKLDIGTDFLDYLAIGAFPAFDSVC TLQVLNODETPLLISLVFGGVYNDATSVVFNAIQSFDLTHLNEAAFHLLGNFLYL FLLSLGAATGLISAYIKKLYPGRHSTREVALMMLMAYLSYMAELDFLGSILTY FFCGIVMSHYTHNVTESRITTKHTFATISFLAETIFLYVGMADLIDIKWRSVSDT PGTSAIVSSILMGVAVGAAFPPLSFLSLAKKNOSEINFNMQVVIWWSGLMRGA VSMALAYNKFTFRAGTIDVRGNAMITSTITVCLFSTVVFGLMTKPLISYLLPHONATT SMLSDNTPKSIHIPLLDQSFTEPPSNHNVPDPDSIRGLTRPRTVHYWYRQFDDS EMRPVFGRGFVPGSPTERNPPDLKA"
BASE COUNT 388 a 346 c 358 g 522 t
ORIGIN
Query Match 74.1%; Score 1614; DB 12; Length 1614;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1614; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 286 atgttggattctctagtgtcgaaactgccttcgttatcgacatctgatcacgcttctgtg 345
Db 1 ATGTTGGATTCTAGTGTGCAAACTGCCTTCGTTATCGACATCTGATCAGCTTCCTGTG 60
Qy 346 gtccgttgaactcctttgttgcactctttgtgcttgattgttctgttgcactctttg 405
Db 61 GTTGGTTGAATCTCTTTGTGCACCTCTTTGTGCTTGTATTGTTCTTGGTCACTCTTTG 120
Qy 406 gaagagaatagatggatgaacgaatccatcacccgcttctgttgcattgggctaggcactggt 465
Db 121 GAAGAGAATAGATGATGAGAACGAATCCATCACCGCCTTGTGATTTGGCTAGGCACCTGGT 180
Qy 466 gttaccattttgttgattagtagtaaggaaagctcgcacattctctctttagtgaagat 525
Db 181 GTTACCACTTTTGTGATTAGTAAAGAAAGAAAGCTCGCATCTCTCGTCTTTAGTGAAGAT 240
Qy 526 cttttctcatatatcttttggccaccattatattcaatcagggttttaagtaaaaaag 585
Db 241 CTTTCTTCATATATCTTTTGGCCACCCCATATATATCAATGCAGGGTTTCAAGTAAAAAAG 300
Qy 586 aagcagtttttccgcaatttctgtagactattatgctttttgtgctgttgggactattatt 645
Db 301 AAGCAGTTTTTCCGCAATTTTCGTGACTATTATGCTTTTTTGGTCTGTGTGGACTATTATT 360
Qy 646 tcttgacaatcatatctcttaggtgtaaacacagttcttctaagaagttggacattggaacc 705
Db 361 TCATTGCAAAATCATATCTCTAGGTGTAAACAGTCTCTTTAAGAAGTTGGACATTGGAACC 420
Qy 706 tttagcttgggtgattatctctgtctattgttgccatatatttctgccaacagattcagtagt 765
Db 421 TTTGACTTGGGTGATATCTTCTATTGGTGCCATATATTGCTGCAACAGATTCAGTATGT 480
Qy 766 aactgcaggtctctgaatcaagacgagacacaccttctttacagttctgttattcggagag 825

Db	481		ACACTGCGAGTTCTGAATCAACGACGACACCTTGGCTTTACAGTCTGTGATTCGGAGAG	540
QY	826	ggtgttgtaagtatgcaacgtcagttggtgttcttcaacgcgatacagagctttgatctc	885	
Db	541	GGTGTGTGAATGATGATCAACGTCAGTTGTGGTCTTCAACGGGATTCAGAGCTTTGATCTC	600	
QY	886	actacacctaaacacgaagctgcttttcatcttcttggaaactcttgattttctc	945	
Db	601	ACTCACCTAAACACGAAGCTGCTTTTCATCTTCTGTGGAACACTTCTTGATTTGTTC	660	
QY	946	ctaagtaaccttctgttgcgtcaacccggtctgataagtcggtatgtttatcaagaagcta	1005	
Db	661	CYAAGTACCTTCTCTGGTGCTCAACCGCTGTGATAAGTCGTATGTTATCAAGAAGCTA	720	
QY	1006	tactttggaagcactcaactgacccgagaggttgccttatgatgcttatggcgatatctt	1065	
Db	721	TACTTTGGAAGGCACTCAACTGACCGGAGAGGTGGCCCTTATGATGCTTTATGGCGTATCTT	780	
QY	1066	tcttatagcttgcagcttttcaacttgacgtatccctcaactgatttttcttctggt	1125	
Db	781	TCTTATATGCTTGTGAGCTTTTTCGACTTGGACGGTATCCTCACTGTGTTTTTCTGTGGT	840	
QY	1126	attgtgatctccattacacatggcacatgttaacggagagctcaagaatacaacaagaag	1185	
Db	841	ATGTGTGATGTCCTCCATTACACATGGCACAAATGTAACGGAGAGCTCAAGAATAACAACAAG	900	
QY	1186	catacetttgcgaacttctgcttcttcttgcggagacatttatttcttctgattggaatg	1245	
Db	901	CATACCTTTGCAACTTTGTTCATTTCTTGGGAGACATTTATTTCTTGTATGTTGGAAATG	960	
QY	1246	gatgccttgacatgacaagtggaagccgtgagtgacacacccgggaacatcgatcgca	1305	
Db	961	GATGCCCTTGGACATTGACAAGTGGAGATCCGTGAGTGACACACCGGGACATCGATCCGA	1020	
QY	1306	gtgagctcaatcctaaatgggtctggtctatggttggaaagcagcgttctgttcttcgcta	1365	
Db	1021	GTGAGCTCAATCCTAATGGTCTGGTCTATGTTGGAAGAGCAGCGTTCGCTTTCCGTTA	1080	
QY	1366	tcgttctatcctaacttaacgaagaatacaaacgagagaatacaacttaaacatgcag	1425	
Db	1081	TCGTTTCTATCTAACTTAGCCAAAGAAGATCAAGCGGAGAAATCACTTTAACATGCAG	1140	
QY	1426	gttgtgatttgggtctggtctcatgagaggtgctgtatctatgctcttgcttgcatacaac	1485	
Db	1141	GTGTGATTTGGTGTCTGCTCATGAGAGTGCTGTATCTATGCTTTCATACAAC	1200	
QY	1486	aagttacaagggccgggacacagatgtacgcgggaatgcgaatcatgatacagagtaag	1545	
Db	1201	AAGTTTACAAGGGCCGGCACAGATGTACGCGGGAATGCAATCATGATCATCAGAGTACG	1260	
QY	1546	ataactgtctgtcttttagcacagtggtgttggatgtatgctaccaaaccactcataagc	1605	
Db	1261	ATAACTGTCTGTCTTTTATGACAGTGGTGTGTTGGTATGCTGACCAACCACTATAAGC	1320	
QY	1606	tacctattaccgacacagacccacacagcatgttatctgatgacacaccccccacaa	1665	
Db	1321	TACCTATTACCGACACGAGCGCCACCGAGCATGTTATCTGATGACAAACCCCAAAA	1380	
QY	1666	tccatacatcccttgttggaccgaagactcgttcaatgagccttcagggaacacacaa	1725	
Db	1381	TCCATACATATCCCTTTGTTGGACCAAGACTCGTTTCATTGAGCCTTCAGGGGAACCAAT	1440	
QY	1726	gtgcctcgccctgacagatatagctggtctcttgacacgcccactcgaaccggtgattac	1785	
Db	1441	GTGCTCGGCCCTGACAGTATACGTGGCTTCTTGACACGGCCCACTCGAACCGTGATTA	1500	
QY	1786	tactggagacaatttgatgactccttcacgcacccgcttcttggaggtcgtgctttgta	1845	
Db	1501	TACTGGAGACAATTGATGACTCCTTTCATCGACCGCTCTTTGGAGGTGCTGGCTTTGTA	1560	
QY	1846	cccttgttccaggttctcacaactgagagaacccctcctgatcttagtaagct	1899	
Db	1561	CCCTTTGTTCCAGTTCTCCAACTGAGAGAAACCCCTCTGATCTTAGTAGAGGCT	1614	
RESULT	3			
LOCUS	AB038492			
DEFINITION	AB038492	2637 bp	mRNA	PLN
	Atriplex gmelini AgNHX1 mRNA for Na/H antiporter Nhx1, complete			18-AUG-2000
	cds.			
ACCESSION	AB038492			
VERSION	AB038492.1	GI:9857313		
KEYWORDS	Na/H antiporter Nhx1.			
SOURCE	Atriplex gmelini cDNA to mRNA.			
ORGANISM	Atriplex gmelini			
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
	Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;			
	Caryophyllales; Chenopodiaceae; Atriplex.			
REFERENCE	1 (sites)			
AUTHORS	Hamada, A., Hayakawa, T., Shono, M., Xia, T., Hayashi, Y. and Tanaka, A.			
TITLE	Isolation and characterization of a Na/H antiporter gene from the			
	halophyte Atriplex gmelini			
JOURNAL	Unpublished (2000)			
REFERENCE	2 (bases 1 to 2637)			
AUTHORS	Hayakawa, T., Hamada, A. and Tanaka, A.			
TITLE	Direct Submission			
JOURNAL	Submitted (17-FEB-2000) to the DDBJ/EMBL/GenBank databases.			
	Takahiko Hayakawa, Plantech Research Institute, Research Center;			
	1000 Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan			
	(E-mail: pri0012@cc.m.kagaku.co.jp, Tel: +81-45-963-3520,			
	Fax: +81-45-962-7492)			
FEATURES	Location/Qualifiers			
source	1. .2637			
	/organism="Atriplex gmelini"			
gene	/db_xref="taxon:118074"			
	682. .2349			
CDS	/gene="AgNHX1"			
	682. .2349			
	/gene="AgNHX1"			
	/codon_start=1			
	/product="Na/H antiporter Nhx1"			
	/protein_id="BAB11940.1"			
	/db_xref="GI:9857313"			
	/translation="MWSQLSSLLSGKMDALTSDHASVVSVMNLFVALLCGCIVIGHLL			
	ENRNWNESTALLIGLATGVVILSGKSSHLVFSDFLFVILPPIIENAGFOV			
	KKQFFNFITIVLFGAVGLVSFTIISLGALSIFKLLDIGLEADLYAIGAFANT			
	DSVCTLOVLDQDETPLYLSVFGVEGNDATSVVFNALQSFDLRIDIRIALQFMGN			
	FLYLFATSLILGATGLLSAYIIKLYFGHSTDRVALMMLAYLSYMLAEFLYLSG			
	ILTVFFCGIVMSHYTWHNTESRVTTHKAFATLSFAEVFLFLYVMDALDIEKWR			
	VSDPGISVAVSSILLGLVMGRAPVFLSWLMNFAKKSQSEKVTFNQOIVIMWAGL			
	MRCVSMALAYNQFTBSGHTOLRGNAIMLTSTISVVLFTSMVFGLLTKPLIMFLPOP			
	KHFTSCSTVSDGSPKSYSLPLLEGNQDYEDVGNHEDTTEPTIVRPSSILRMLLN			
	APTHTVHHYWRKFDDFSFRPVFGGFGVFPVFGSPTEQSTNNLVDRT"			
	2607			
polyA_site	/notes="30 a nucleotides"			
BASE COUNT	685 a 509 c 530 g 913 t			
ORIGIN				
	Query Match 38.2%; Score 832.8; DB 12; Length 2637;			
	Best Local Similarity 71.9%; Pred. No. 7.1e-175;			
	Matches 1146; Conservative 0; Mismatches 412; Indels 36; Gaps 3;			
QY	323	cgacatctgacacagcttctgtggttgctgaatctctttgttgcaactcttctgtgctt	382	
Db	731	CCACTTCTGATCAGCGCTTCTGTGTCTCGATGAACCTTTGTGTGGCACTGTTATGTGGTT	790	
QY	383	gtattgtcttggctcatcttttggaaagaaatagatggaacgaatccatcacccgct	442	
Db	791	GTATCGTAATTTGGTCATCTTCTAGAGGAGATCGTGTGATGAATGAGTCCATCACTGCC	850	
QY	443	tgttgattgggctaggcactggtgttaccatttctgttattagtagtaagggaagacgcgc	502	
Db	851	TTCCTTATAGTTGGCTACTGGGGTGTGTGATCTGCTGATTTCTGATGGGAGAAAAGTTTAC	910	


```
source 1...2313
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
297...1904
/gene="OsNHX1"
297...1904
CDS
/gene="OsNHX1"
/note="similar to yeast sodium/proton exchanger"
/protein_id="BA083337.1"
/db_xref="GI:5731737"
/translation="MGMEVAARLALGALYTTSDYASVSVINFLVALLCACIVLGLLEE
NRWNESITAIIGLCVGVVILLMTKSGSHLFSEDFLYLLPPIFNAGFQVK
KOFFRPMWITLFGAVGTMSFFETIAAIAFVSNMIGTLDVGDFAIGAIFSATDS
VCTQLVNMDDTPEPLSLVPEGVNDASIVLENALONFDLHVHDAAVLKLPLGNF
YLFSSPLGVFAGLLSAYIIKKLYICRHSOTREVALMLMAYLSYMLAEILLDLSGIL
YFPCGIVMSHYTHWNTESRVTTHAFATLSFAETFELELYVGMOKLDIEKWEFAS
DPRKSGISILGLVIGRAAFVPLSLNFKAPNEKITWRQOVVFWAGLMR
GAVSLAIYNKFTKSGHTQJHNAIMITSTIVLFTVFMVGMFTKPLIRLLPASGH
PVTSEPSKSLHSPLLTSMOGDLESTTNIVRPSSRLMLTKPTHVHYWRKEDDA
LMRPMFGRGFVPPSPGSPTEQSHGR"
2313
polyA_site
/note="17 a nucleotides"
BASE COUNT 543 a 538 c 569 g 563 t
ORIGIN

Query Match 35.3%; Score 768.8; DB 12; Length 2313;
Best Local Similarity 68.6%; Pred. No. 1.2e-160;
Matches 1077; Conservative 0; Mismatches 487; Indels 6; Gaps 1;

QY 312 qccttcgttaccgacatcgtacacgcttctgtgttgcttgaatctcttctgtcact 371
DB 329 GGGCTGTGTACACGACCTCCGACGTACGCGTGGTGTCTCATACACCTGTCTCGCGCT 388
QY 372 tcttctgttctgttctgttctgttctgttctgttctgttctgttctgttctgtt 431
DB 389 GCTCTGCGCTGCATGCTCTCGGCACCTCTCGAGGAGAAATCGCTGGGTCAATGAGTC 448
QY 432 catcacgccttctgttctgttctgttctgttctgttctgttctgttctgttctgtt 491
DB 449 CATCACGCGCTCATCATCGGCTCTGCACCGCGCTGTGTGTGTGTGTGTGTGTGTGT 508
QY 492 aaaaagctcgcattctctgttctgttctgttctgttctgttctgttctgttctgttct 551
DB 509 GAAGAGCTCGCACTTATTCGTCTCAGTGAGGATCTCTTCTCATCTTACCTTCCTCCTCC 568
QY 552 cattattatcaatgcagggttcaagtaaaagaaagcagtttctccgcgaattctgtgac 611
DB 569 GATCATCTTCAATGCAGAGTTCAGGTAAAGAAAAGCAATCTTCCGGAAATTTTCATGAC 628
QY 612 tattatgtcttctgttctgttctgttctgttctgttctgttctgttctgttctgttctgt 671
DB 629 GATCATATTTTGGAGCGCTCGGACCAATGATATCTTTTTCACAAATATCTATTTCTGTC 688
QY 672 aacacagttcttcaagaagttggaattggaaccttctgttctgttctgttctgttctgtat 731
DB 689 CATTTGCAATATTCAGCAGAATGAACATTTGGAACGCTGATGTAGGAGATTTTCTTCAAT 748
QY 732 tgggtgcattattgtgcaacagattcagatgttacactgcaggttctgaatcaagaaga 791
DB 749 TGGAGGCACTTTTCTGCGACAGATTTCTGTCTGCACATTCGAGGTCTCTCAATCAGGATGA 808
QY 792 gacaccttcttctacagcttctgttctgttctgttctgttctgttctgttctgttctgt 851
DB 809 GACACCTTTTGTACAGTCTGTATTCTGGTGAAGGTGTTGTGAACGATGCTACATCAAT 868
QY 852 tgggtcttcaacgagattcagagtttctgttctgttctgttctgttctgttctgttctgt 911
DB 869 TGTCTTTTCAACGCACTACAGAACTTTGATCTTGTCTGACATAGATCGGCTCTGCTTCT 928
QY 912 tcatcttcttggaaactcttctgttctgttctgttctgttctgttctgttctgttctgt 971
```

```
DB 929 GAAATTCCTTGGGAACTTCTTTTATTTATTTTGTGAGCACCTTCTTGGAGTATTGCG 988
QY 972 cgtctgataagtcgtatgttatacaagaagctatactcttgggaagcactcaactgacgc 1031
DB 989 TGGATTGCTCAGTCATACATAATCAAGAAGCTATACATTGGAAGCA/TTCTACTGACCG 1048
QY 1032 aagagttcccttctgtatgtatgagcgtatcttcttcttcttctgttctgttctgttctgt 1091
DB 1049 TGAGGTGCGCTTATGATGCTCATGCTTACCTTTTATATATGCTGCTGAGTGTCTAGA 1108
QY 1092 cttgagcggatctccctcactcgtgttcttctgtgtgtatgtcaccattacacatggca 1151
DB 1109 TTTGAGCGCATCTCCCGCTATTCTTCTGTGTGTTATGTAATGTCACATTACACTTGGCA 1168
QY 1152 caatgttaacggagagctcaagaatacaacaagaacacacaccttggcaacttctgttcttct 1211
DB 1169 TACGCTCAGAGAGATTCAAGAGTTACAACAAAGCAGCATTTGGCAACTCTCTCTCTTCT 1228
QY 1212 tgcggagacatttatttcttctgtatgttggatgtgaccttggacatggacaagtggag 1271
DB 1229 TGTGAGACTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1288
QY 1272 atccatgagtaacacacacgggaacatcgatcgagcagtcgaatcaatcaatggctcgt 1331
DB 1289 GTTTGCCAGTGACACCTGCCAATTCATTTGGATTAAGCTCAATTTTCTGCTAGGATTGT 1348
QY 1332 catggttgaagcagcgttctgttcttctgttctgttctgttctgttctgttctgttctgt 1391
DB 1349 TCTGATTGGAAGAGCTGCTTTTGTATTCCCGCTGTCTTCTTCTGCAACTTACAAAGAA 1408
QY 1392 gaatcaacggagagaaatcaacttaacatgcaggttctgttctgttctgttctgttctgt 1451
DB 1409 GGCACCGAATGAATAATAACCTGGAGACAGCAAGTTGTAATATGTTGGCTGGGCTGAT 1468
QY 1452 gagagtgctgtatctatgttcttctgttctgttctgttctgttctgttctgttctgttctgt 1511
DB 1469 GAGAGAGCTGTGTGATTTGCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 1528
QY 1512 tgcacggggaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatca 1571
DB 1529 GCTGACGCGCAATGAATAATGATCACCAGCACCACATCCTGCTGCTTCTTCTTCTTCTTCT 1588
QY 1572 ggtgttctgtatgtgcacaaacacacacacacacacacacacacacacacacacacacac 1631
DB 1589 GGTATTGGATGATGACAAAGCCATTTGATCAGGCTGCTGCTACCGGCTTCAGGCCATTCCT 1648
QY 1632 cagcagcatgttatctgtatgacacacacacacacacacacacacacacacacacacacac 1691
DB 1649 TGTACCC-----TCTGAGCCTTTCATCACCAGTCCCTGCTGCTTCTTCTTCTTCTTCT 1702
QY 1692 agactcgttctgtatgtcaggaacacacacacacacacacacacacacacacacacacacac 1751
DB 1703 CATGCAAGTTTCTGACCTCGAGAGTACAAACCAACATTTGTGAGGCTTTCAGGCCCTCCGGAT 1762
QY 1752 cttcttgacacggccactcgaacgcgtgcttactactgagacacacacacacacacacacac 1811
DB 1763 GCTCTTCAACAGCGCCACACTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1822
QY 1812 catgcacacgcttctgtatgtcgttctgttctgttctgttctgttctgttctgttctgttct 1871
DB 1823 GATGCGACCGATGTTTGGCGGCGCGGCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 1882
QY 1872 gagaaacacct 1881
DB 1883 GCAGAGCCAT 1892
```

```
RESULT 6
TM021B04/c 90019 bp DNA PLN 12-JUN-1997
LOCUS Arabidopsis thaliana BAC TM021B04.
DEFINITION AF007271
ACCESSION
```



```
LSQRDGLLTQNGNMSKSPETRPVSSCKTKENTFENQORSWQFQIDQIWAII
SNDKSPKRYAQIKKIDTSPFKLHVAPLELYRDEIHMPRVCCGCRFKLTGAEEVV
PSFSHOWKAVKATKKRREVEYVPGKEIWALYKNCNTRDYETELEIVEVETDEORI
QAMTLTAKGFNNKPLRYRSEBSNASFINERNLCBEDSDSDIMLLMVGVSDAYESQI
PRMRKQITLQKGAHTIQALKDADDPHFQRLYRMMPEVEFAELCHLLQMKTLGUKTPHVC
VEEMVATLITVQNSRYCHYDFTFKRSKFSSTINFKVLRALNMLATLMAKYNTNV
PSIKSTTRFPYFDCVQADIGDTHINAMVQPEKASYNRKGVISQNVLAACNFDE
FYIULSGWEGSAHSDVLQDALTRTNRLQVPEGNRLQVLKGTMGSDPKPKRKGDY
NPMSPETKLLVOLLVEGINNNWRDSNGTISKLTVEFKFPEINKERCSKNYHYS
RMKYLKIQOCLDLQRFSSGFGDPDLTKRETASDEWSDYLKAKHPNNKQLRYDTFEE
FDELQIIFEGVATKNAIGLCSDDGLTYRAGENPRKEYVDDDFNDVEYDTHHES
SHYAPFHWGTSSEPKLPKRRTSRSESTSOKEPSMPVSSKILLDIOQREERQON
EYAAKKNVMDAIEKISDLDECIRYKALTKIYHLGIQDFVFSMSVEERLGIQIME"
complement(34437. 37759)
/genes="A.TM021B04. 8"
complement(join(34437. 35003,35094. 35272,35446. 36036,
36081. 36197,36313. 36607,36724. 36905,37102. 37759))
/genes="A.TM021B04. 8"
/notes="contain similarity to type 1 inositol
1,4,5-triphosphate receptors"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB61078.1"
/db_xref="GI:2191193"
/translation="MKAKQSEFEKKEKDFLEQKAEVEKRRREVEQLEKFTTRMESVE
RVSDEKLMELGURATELEKMEVEKHEKRIYAGDKLGRPEPLVSLAKNGLSVTM
PVKCSLVLNENADEMKNTALARMVPLDPAKVLVLDALIEGSEKYEYWKDLGEADDR
VYNSVILLENLKNLKITPOVQOEATPLGIWNLGKAKANKNDPPOVFOCALFLAA
YGAATLKKEBIELAYLICEFRULIKLCPGPRGELLIEFFSDSKAARYIAGTGS
MDQAARKREKKADAMAIFYKEAKETMPAKILKRLAVKNDSEARAMEPVQS
YERQSTGTFGVEKSEAKSIPYEQKAEVTKRPLTEPTAPSONLTVCFNMKHELTYSV
NLSIFVGETLRHKNONIFLLFKILKQPEVVCPTQVKVESGADHOPDTIATHPSG
TETKLNILSGSTKADMLRELVEKOPKLESDLSNALKCTPDPAKFLDTSNALCPTNT
EGGYEFKMLIISGSSCLLNLQKLLPKLPIHPVKGDAKKLAVYWKDKIAKSKROLEV
ICFLQFLGFIYSVSESLLDLGLDNLNSYQWTVSPDICFLGLDDAIPGFIQNLKIG
HRIKAIDIYSFQVHVRFPQSAIINDSLRTIKESAESYREAKNESTQVAAIDROV
RALRAAIKICISCHLESEFQGLDEEQIKSLKLLKRLNRTSNGSGSSASRKPDPITKQS
QTAKPPTVAEYAPVTSNPLPSTEAASSASAKPFSKKNKRGKRSNGNNOSSGHIA
SHTSNHYPSHDYSLNQLRTWPVDNYDRGTFGFPNPDYNNQWQGPQGFYHLYQLPD
PRYRN"
complement(39523. 43518)
/genes="A.TM021B04. 7"
complement(join(39523. 39921,40013. 40173,40283. 40858,
41109. 43518))
/genes="A.TM021B04. 7"
/notes="contains similarity to tropomyosin and kinesin"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB61071.1"
/db_xref="GI:2191186"
/translation="NEEIKLENIIRLCDVKANNIRKTMWMIKSQASDVLILNLQWDF
```

```
Query Match 31.8%; Score 693; DB 15; Length 90019;
Best Local Similarity 69.1%; Pred. No. 1.le-143;
Matches 1379; Conservative 0; Mismatches 0; Indels 616; Gaps 7;

QY 773 aggttctgaatcaagcagacaccttgccttacagctcttgcagagaggtgttg 832
|||||
Db 66754 AGGTTCTGAATCAAGACGAGACACCTTGTCTTACAGTCTTGTATTCGGAGGGGTG 66695
|||||

QY 833 tgaatgatgcaacgctgagttggtcttcaacgcgattcagagcttgcattcaactcacc 892
|||||
Db 66694 TGAATGATCAACGTCAGTTGTGTCCTTCAACGCGATTTCAGAGCTTGTATCTCACTCACC 66635
|||||

QY 893 taaccacgaagctgcttttcattctcttgcgaactcttgattgtttcttctaagta 952
|||||
Db 66634 TAAACCACGAAGCTGCTTTTCATCTCTTGGAACCTCTCTGTATTTGTTCTCTCAAGTA 66575
|||||

QY 953 ccttgcttggtgctgca----- 969
|||||

Db 66574 CCTTGCCTGGTGCTGCAGTAAGTCCATTTCTTCCTGCCATTCCTTCTTGTAGCTG 66515
|||||

QY 970 -----accggtcgtgataag 983
|||||
```

```
Db 66514 CTATATGACACGCTCTGGTGTTTCACATCTTTTGTATTTCATCTTGCAGACCGGTCTGATAAG 66455
|||||

QY 984 tgcgtatgttatacaagaagctatacttttgg----- 1014
|||||

Db 66454 TCGCTATGTTATCAAGAAGCTATACCTTTTGGGAAGGTACATATTTAACGTCATATATATGAC 66395
|||||

QY 1015 ----- 1014
|||||

Db 66394 TTGTTTCTTTAAGTTGTTTCTGTTGGTCTAACGTAGATAGTGGCCTCACATTAATTATGTCA 66335
|||||

QY 1015 -aggcaactcaactgaccgagagaggttgccttattgatcgttatggcgtatctttcttatat 1073
|||||

Db 66334 CAGGCACCTCAACTGACCGAGAGGTTGCCCTTATGATGCTTATGCGGTATCTTCTCTTATAT 66275
|||||

QY 1074 gcttgctg----- 1081
|||||

Db 66274 GCTTGTCTGAGGTTAGATTCTTTTAAATAACACGAACATATTTTCATTTCAGTTGCATTTCAAC 66215
|||||

QY 1082 -----agcttttcgacttgagcgggtatcctc 1107
|||||

Db 66214 ACAAAATTATATGATGTTATATTAATTTGGTGTGCGACTTTTCGACTTTCGCGGTATCCTC 66155
|||||

QY 1108 actgtgttttctctggttattgtgatgtcccatcacacatgacacaatgtaacggagagc 1167
|||||

Db 66154 ACTGTGTTTTCTGTGGTATTGTGATGTGCCATTACACATGCGCAATGTAACGAGAGC 66095
|||||

QY 1168 tcaagaatacaacacaaa----- 1184
|||||

Db 66094 TCAAGAATAACAAACAAAGTATGCAAACTTTAGTTTCTCCGCTTTTGTGCGAGATTAACAC 66035
|||||

QY 1185 -----gcatacctttgcaa 1198
|||||

Db 66034 TTTCTACCTCCCAATGATTATCTTATGATTTTTCGATGAACAGGCATACCTTTGCAG 65975
|||||

QY 1199 ctttgtcattcttcgagagacatttatttcttctgtatgttggaatggagtcgcttgagca 1258
|||||

Db 65974 CTTTGTCTATCTTTCGCGGAGACATTTATTTTCTTGTATGTTGGAATGATGCGCTTGAGCA 65915
|||||

QY 1259 ttgacaagtggagatccgtgagtgacacacccgggaaacatcgatcgagtgagctcaatcc 1318
|||||

Db 65914 TTGCAAGTGGAGATTCCTGTGAGTGACACACCGGGAACATCGATCGAGTGAAGTCAATCC 65855
|||||

QY 1319 taatgggtctggtcatggttggaagagcagcgttcgctcttccttcgttatcgctttctatcta 1378
|||||

Db 65854 TAATGGGHTCTGCTCATGTTTGGAAAGCAGCGCTTCGCTTTCCGTTATCGTTTCTTCTACTA 65795
|||||

QY 1379 acttagccaagaagaatcaagaagcagagaaaaatcaactttaacatgc----- 1423
|||||

Db 65794 ACTTAGCCAAGAAGAAATCAAAAGCGAGAAAAATCAACTTTAACATGCAGTATGTATCATTC 65735
|||||

QY 1424 ----- 1423
|||||

Db 65734 TACTTGTACTAATAGATATCAATAATTTCTCTGGATTTTCAGCTATTTATATATCTATCATTT 65675
|||||

QY 1424 -----aggtttgatcttggtctggtcttcctcatgagagtgctgctgattcatg 1470
|||||

Db 65674 TTTCCCATTTCTTAGGTTGTGATTTTGGTGTCTGTCTCATGAGAGGTGCTGTATCATG 65615
|||||

QY 1471 gctcttgatacaaca----- 1486
|||||

Db 65614 GCTCTGTGCATACACAAGGTATATAAAGAAATGCTCTTTAAATACTCTCTATGAAGCTTCAA 65555
|||||

QY 1487 -----agttacaagggcgcgagacaca 1509
|||||

Db 65554 ATGACAAAATCATCATTTAAACACTTGTCTTCTATTGTAGTTTACAAAGGGCGGCGACACA 65495
|||||

QY 1510 gatgtacgggaatgaatcaatcatgatcacgagtacgataaactgctgctcttttttagaca 1569
|||||

Db 65494 GATGTACGGGGAATGCAATCATGATCACGAGTACGATAACTGTCTGCTTTTTTAGCACA 65435
|||||

QY 1570 gt----- 1571
|||||
```

Db 65434 GTGGTAAGAAATCTCTGGATGATCAAAATGAGTTACATAAATATTTCTTAATTGAATTTGGAC 65375
QY 1572 -----gggtgttggtatgctgaccaaacc 1596
Db 65374 ATTTGGTTAATAGATTCTTTGAACCTTGTTTTCAGGTGTTTGGTATGCTGACCAAAACCA 65315
QY 1597 ctcataaactaactattaccgcacacagacccacccacagacatggttatctgtatgacaaac 1656
Db 65314 CTCATAAGCTACCTATTACCACACCAAGAACGCCACCGAGCATGTTATCTGATGACAAC 65255
QY 1657 accccaaaatccatacatatccctttgttgacaaacactcgttcattgaccttcagg 1716
Db 65254 ACCCCAAATCCATACATATCCCTTTGTTGGACCAAGACTCGTTCAATGAGCCPTTCAGGG 65195
QY 1717 aaccacaatgctcgcctgcagcagatatacgtggtctcttgacacgcccacactcgaacc 1776
Db 65194 AACCAATATGTCCTCGCCCTGACAGTATAGTGGCTCTTTGACACGGCCACTCGAACC 65135
QY 1777 gtgcattactactgagacaatttgatgactccttcattgaccccgctctttggaggtcgt 1836
Db 65134 GTGCATTACTACTGGAGACAATTTGATGACTCCTTCATGCGACCCGCTCTTTGGAGGTCTGT 65075
QY 1837 ggccttgaccctttgttccaggtctccaaactgagagaaacccctcctgattagtaag 1896
Db 65074 GGCCTTGTACCCCTTTGTTCCAGGTTCTCCAACTGAGAGAAACCCCTCCTGATCTTAGTAAG 65015
QY 1897 gcttgagggtgaacgtggaagaaacgttctgatttttttggtagaaggggtgattcaaa 1956
Db 65014 GCTTGAGGGTAACGTGGAAGAAAGCTTTGATTTTTTGGTAGAAAGGGTGATTCAAA 64955
QY 1957 ttatgctttgtgtaaatatccattgtgaatatgtttgttgagagacagaaatcgtcct 2016
Db 64954 TTATGCTTTTGTGTAAATATATCCATTTTGTAAATATTTGTTGAGGACAGAAATCTGTCT 64895
QY 2017 aacgttttgagacagaagaacaaacatgcaacttgaaacttgaaactgttattgattatgta 2076
Db 64894 AACGTTTGGAGACGAAGAAGCAAAACATGGCACTTTGAAAGTGTGTTGATGATGATGTA 64835
QY 2077 attatattcatattgtttgttgaacacaaactacacattgtttatgttttgaattt 2136
Db 64834 ATTATATTCAATATTGTTTGTGTGAACACAAACTACACATTTCTTTATGTTTGAATTT 64775
QY 2137 ggttttgccttcogaa 2151
Db 64774 GGTTTTTCGCTTCGAA 64760

RESULT 7
AY028416 600 bp mRNA PLN 02-APR-2001
LOCUS Citrus x paradisi sodium/proton exchanger mRNA, partial cds.
DEFINITION Citrus x paradisi
ACCESSION AY028416
VERSION AY028416.1 GI:13508486
KEYWORDS
SOURCE Citrus x paradisi.
ORGANISM Citrus x paradisi
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCE 1 (bases 1 to 600)
AUTHORS Porat,R., Lurie,S. and Povancello,D.
TITLE A heat treatment induced the transcription of a sodium proton exchanger gene in grapefruit peel tissue
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 600)
AUTHORS Porat,R., Lurie,S. and Povancello,D.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2001) Postharvest Sciences, ARO, The Volcani Center, Bet Dagan 50250, Israel
FEATURES
source 1..600
/organism="Citrus x paradisi"
/db_xref="taxon:37656"

CDS
/tissue_type="peel"
/note="heat-induced"
132..>600
/note="ion transporter"
/codon_start=1
/product="sodium/proton exchanger"
/protein_id="AAK27314.1"
/db_xref="GI:13508487"
/translation="MDQAISSVVRKLQMVNTSDHNSVVSIINFVALPCASIVIGHLLE
ESRWNESITALLIGVCAGVILLTGTGKSSHLFVSEDLFFIVLPPIFNAGFOVK
KKQFFNFETIMFGAIGTLYSCTIISLGVIQFFKKLDIGTLDIGDYLGAIGAF"
BASE COUNT 170 a 86 c 141 g 203 t
ORIGIN

Query Match 12.2%; Score 266.6; DB 14; Length 600;
Best Local Similarity 72.1%; Pred. No. 4, 4e-49;
Matches 347; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
QY 265 tggatctagaagaagataacaatgttggtatctctagtcgaaaactgccttcggttatcg 324
Db 120 TGGTTGCAAAAAATGGATCAAGCAATAAGTTCTGTGTTAAGGAATTTGCAAAATGGTGAAC 179
QY 325 acatctgatcaagcttctgtggttgcgttggaatctctctttgttgcaactctctttgtgtcctt 384
Db 180 ACCTCTGATCACAACTCGGTGGTTTCCATAAAACATTTTTTGTGGCGCTTCTCTGTGCTTCT 239
QY 385 attgtcttggctcatcttttggaaagagaatagatgatgaacgaatccatcaccgcttg 444
Db 240 ATTGTGATTTGGTGCACCTCTTTGAAGAAACCCGGTGATGAACGAGTCCATTTACAGCTCTA 299
QY 445 ttgattgggtcaggcactggtgttaccattttgttgattagtaaaggaaaaagcgcgat 504
Db 300 CTGATTGGAGTTTGTGCAGGTGTGATTAATTTGCTGCACAACTGGGGGAAAAAGCTCACAT 359
QY 505 ctctcgtctttagtgaagatctttcttcacatctctctctctctctctctctctctctctctct 564
Db 360 CTGTTTGTCTCAGTGAAGATCTGTTCTTTATATATGTTGCCCCCATTAATATTAAT 419
QY 565 gcagggttccaagtaaaaaagacaggttttccgcaatttcgtgactattatgctttt 624
Db 420 GCAGGATTTCAGGTGAAAAGAACAAATTTTCCGTAACCTTACTATCATCATGCTGTT 479
QY 625 ggtcgtgttgggactattattcttctgcacaatctatctctctctctctctctctctctctct 684
Db 480 GGTGCAATTTGGGACATTAGTAAGCTGTACCATCATCATCACTAGGTGTTATTTCAGTCTTT 539
QY 685 aagaattggacattggaaccttgacttgggtggtgattatcttctgatttggtgcacattt 744
Db 540 AAGAAATTTGGATTTGGTACTCTGGATATAGGAGATTATCTAGCAATTTGGTGAATATTT 599
QY 745 g 745
Db 600 G 600
RESULT 8
AF279671 421 bp mRNA PLN 02-AUG-2000
LOCUS Mesembryanthemum crystallinum putative Na+/H+-antiporter (NHX2)
DEFINITION mRNA, partial cds.
ACCESSION AF279671
VERSION AF279671.1 GI:9652169
KEYWORDS
SOURCE common ice plant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae; Caryophylliales; Alzooaceae; Mesembryanthemum.
REFERENCE 1 (bases 1 to 421)
AUTHORS Chauhan,S., Forsthoefel,N. and Bohnert,H.J.
TITLE Na+/myo-inositol symporters and Na+/H+-antiporters in plant sodium transport

TITLE	Na ⁺ /myo-inositol symporters and Na ⁺ /H ⁺ -antiporters in plant sodium transport
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 459)
AUTHORS	Chauhan,S., Forsthoefel,N. and Bohnert,H.J.
TITLE	Direct Submission
JOURNAL	Submitted (19-JUN-2000) Biochemistry, University of Arizona, 1041 E. Lowell St, Tucson, AZ 85721, USA
FEATURES	Location/Qualifiers
source	1..459
	/organism="Vesembryanthemum crystallinum"
	/db_xref="taxon:3544"
gene	<1..>459
	/gene="NHX1"
CDS	<1..>459
	/gene="NHX1"
	/codon_start=1
	/product="putative Na ⁺ /H ⁺ -antiporter"
	/protein_id="AAF91423.1"
	/db_xref="GI:9652168"
	/translation="pPIIFNAGVQVKKQFFKNSFSLMFGVLGTIISGLISSGVAL LLKKIGLTOLNLDVLAIAIFAISFASDQETPLLYSIVFEGCVNDATSI

	Query Match	9.3%	Score 203.6;	DB 13;	Length 459;	
	Best Local Similarity	65.3%;	Pred. No. 4.3e-35;			
	Matches 299;	Conservative	0;	Mismatches 159;	Indels	0; Gaps 0;
Qy	547	ccaccattatctcaatgcagcggtttccaagtataaaagaacgactgttttcgcgaatttc	606			
Dd	1	CCTCCCAATAACTTCAACGCGGGTTTCAGGTAAAGAAAAGCAGTTTTTCAAGAAGACTTC	60			
Qy	607	gtgacctatcgtcttcttggtgctgctgtgggactattattcttgcacaatcatctctcta	666			
Dd	61	TCAATTCATATTGATGTTTGCGGTACTGGGAACAATAAATCATATTGCGCTCATATCATCA	120			
Qy	667	gggttaacacacaggtctctttaagaagcttgacatcggaacctttgacttgggtgattatctt	726			
Dd	121	GCTGCCGTACTACTGTTGAAGAAGATTGGTTTGACTCAGCTTAATCTTAATGACTATCTG	180			
Qy	727	gctatgtgtgccatatcttgtgcacaagattcagtatcagctgcaggttctgaatcaa	786			
Dd	181	GCCTCTGGTGCCATATTCTCAGCCACAGACTCTGTGTGACTTTGCAGGTCTCTCAATCAG	240			
Qy	787	gacgagacacctcttgcttacagctcttgattccggagaggggtgtgtgatgatgacaaag	846			
Dd	241	GATGAACAACCCTCTCTTTACAGTATTGTATTGTGGAGGGGAGTAGTAATGATCAACC	300			
Qy	847	tcagttgtgtgtctcaacgcgcatcacagcgtttgatctcactcaaccaaccacagaagct	906			
Dd	301	TCCATTGTGCTTTTCAATGCGAGTCCCATCACTTGATTTAAGCAAACCTCAGTTCATGACT	360			
Qy	907	gcttttcactctcttggaaaactcttctgtatttggcttctcctaagctacottgcttggtct	966			
Dd	361	GCTTTAGCATTCGTCGGGAACCTTCTCTACTCTCTTCTCACAGCACAAATCTTGGCATT	420			
Qy	967	gcaacgcgtctcgataagtgcdtatgttatcaagaagct	1004			
Dd	421	CTAGTTGGTTTGTGTGAGTGCTTATGTATTAAAGAAGCT	458			
RESULT	10					
AF307944						
LOCUS	AF307944	320 bp	mRNA	PLN		02-APR-2001
DEFINITION	Zea mays Na+/H+ antiporter mRNA, partial cds.					
ACCESSION	AF307944					
VERSION	AF307944.1	GI:13507162				
KEYWORDS	.					
SOURCE	zeae mays.					
ORGANISM	zeae mays					

NKVS LHAHDISVFF"

NKVS LHA D I S V F F "

RESULT 12
 ATAC011623
 LOCUS
 DEFINITION ATAC011623 86022 bp DNA PLN 24-JAN-2001
 Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence,
 complete sequence.
 AC011623
 VERSION AC011623.5 GI:12408729
 KEYWORDS
 SOURCE HTG.
 ORGANISM
 Arabidopsis thaliana
 thale cress.
 Arabidopsis thaliana
 Magnoliophyta: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Eudicotyledons; core eudicots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 86022)
 Lin.X., Kaul.S., Town.C.D., Benito.M., Creasy.T.H., Haas.B.,
 Bowman.C.M., Koo.H., Fujii.C.Y., Utterback.T.R., Barnstead.M.E.,
 Rongman.C.L., White.O., Nierman.W.C. and Fraser.C.M.
 Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence
 Unpublished
 2 (bases 1 to 86022)
 Lin.X. and Kaul.S.
 Direct Submission
 TITLE Direct Submission
 JOURNAL Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org
 3 (bases 1 to 86022)
 Lin.X.
 Direct Submission
 TITLE Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 On Jan 24, 2001 this sequence version replaced gi:12280803.
 Address all correspondence to:
 Xiaoying Lin
 The Institute for Genomic Research
 9712 Medical Center Dr.
 Rockville, MD 20850, USA
 e-mail: xlin@tigr.org
 BAC clone F24P17 is from Arabidopsis chromosome III and is near the
 molecular marker mi403.
 The orientation of the sequence is from SP6 to T7 end of the BAC
 clone.
 Genes were identified by a combination of three methods: Gene
 prediction programs including GRAIL (available by anonymous ftp
 from arthur.egr.mnrl.gov), GeneFinder (Phil Green, University of
 Washington), Genscan (Chris Burge,
 http://genome.stanford.edu/~chris/GENSCANW.html), and NetPlantGene
 (http://www.cbs.dtu.dk/netgene/cbsnetgene.html), searches of the
 complete sequence against a peptide database and the Arabidopsis
 EST database at TIGR (http://www.tigr.org/tdb/at.html).
 Annotated genes are named to indicate the level of evidence for
 their annotation. Genes with similarity to other proteins are named
 after the database hits. Genes without significant peptide
 similarity but with EST similarity are named as 'unknown' proteins.
 Genes without protein or EST similarity, that are predicted by more
 than two gene prediction programs over most of their length are
 annotated as 'hypothetical' proteins. Genes encoding tRNAs are
 predicted by tRNAscan-SE (Sean Eddy,
 http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
 identified by RepeatMasker (Arian Smit,
 http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
 genomic sequence that are not annotated as genes but have predicted
 exons by GRAIL are annotated as misc features.
 FEATURES
 Location/Qualifiers
 1..86022
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="III"
 /map="mi403"
 /clone="F24P17"
 join(<283..529,638..818,916..1063,1101..1142,1160..1376,
 1511..1822)
 mRNA

gene
 /gene="F24P17.1"
 <283..1822
 /gene="F24P17.1"
 /note="similar to pyruvate dehydrogenase kinase
 GB:AAC97601, [Arabidopsis thaliana]"
 CDS
 join(<283..529,638..818,916..1063,1101..1142,1160..1376,
 1511..1569)
 /gene="F24P17.1"
 /codon_start=1
 /product="putative pyruvate dehydrogenase kinase, 5'
 partial"
 /protein_id="AAF08568.1"
 /db_xref="GI:6437541"
 /translation="VRDWTLESPRMRAPPEIKDSGDEKDFOTMIKAVKVRHNNVPM
 MALCVNOLKKGMNSGNLDEIHOFDRFLYSRIGRIMLGIOHVELHNNPPLHTVGYTH
 TKMSPMVARNASDEARICFREYSGAPEINIGDPSTFPYVPTLHLMYELVKN
 LRAOERFVDSRVAPPIRIIVADGIEDVTIKPFSLLHRDPDIIVVSDGGGIFARSG
 LPRIFTLYSTARNPLEEDVDIGIADVPVTMAGYGYGLPISRLYARFEGDGLQIISME
 GYGTDAVLHLSRLGDSQEPPLP"
 complement(2022..2051)
 /rpt_family="POLY_A"
 /note="exon predicted by xgrail, quality
 excellent_shadowexon"
 repeat_region
 misc_feature
 complement(2053..2372)
 /note="exon predicted by xgrail, quality
 excellent_shadowexon"
 misc_feature
 /note="exon predicted by xgrail, quality excellent"
 2057..2160
 /note="exon predicted by xgrail, quality excellent"
 2287..2363
 /note="exon predicted by xgrail, quality excellent"
 join(<2630..3712,3813..4007,4256..4324,4406..4675,
 4767..5117,5215..5492,5579..56599)
 /gene="F24P17.2"
 <2630..56599
 /gene="F24P17.2"
 /note="similar to RNA helicase (DRH1) GB:BAA28347
 [Arabidopsis thaliana]; Pfam HMM hits: WW/rsp5/WWP domain
 containing proteins, DEAD and DEAH box helicases,
 Helicases conserved C-terminal domain"
 CDS
 join(2630..3712,3813..4007,4256..4324,4406..4675,
 4767..5117,5215..5492,5579..56599)
 /gene="F24P17.2"
 /codon_start=1
 /product="putative RNA helicase"
 /protein_id="AAF08584.1"
 /db_xref="GI:6437557"
 /translation="MATTETPASAGPRVAPEDPTLPQWKGIDGSGTGLYWNPEP
 NVTQYERPSAPPHSATTPKLAOTIPVSSGOGHQAQHPVGHVSQOHGFOOQPOO
 PPSVHPQWQHPQAOQPOQSQOFPQOQSQSMVPHGHPVQVQPTQOQOQOQ
 MNQHSQMPQQLSHQVHNSQHHYGFPHMQTQGLNSHQPOGPHGQOFPQOQOQ
 SLAPRDEDFHGGKTFGSPQHPNRSRSPQNTHEPANAASQTNANLAMAQKNG
 POANAAVTQOQGANLIHQOLGPRAPQMDQTLHQKSHVSPQNNNTYNNLQSRG
 GNDYSYVNRMEVPRGAQPLHPAAMPKDIRISGGPTNADPAMQGTGHTYGHAGAP
 PKSLVPRPHFTSPDVPPLSPVETIYRKHQEVTTGTENIPAPYITFESSGLPELREL
 LSGAGFPPTQAQWPIALQSRDIAVAITGSGTKTGLYLIPAFILHRCDNRNRP
 TVLLAPTRLATQIDAEALFGRSSRISCTCLYGAQPKPOLKELEGADIVATPG
 RLNDLEMKMTDFQVSLVLDADRLMDGFEPOIRKIVNEIPPRROTLMYATWPK
 EVRKTSDDLVPQVNIQGVDELAANKAITQYVEVPMKERLEOILRSQERGSK
 VIIICSRKLCDLHARSYGRFGAVYIHGDKTQGRDMLNLFQPSKSLIATDVA
 RGLDKLRLVYINIDFTPTGVEDYVIRHIGTGRAGATGVAFTEFTQDKKYPADLIKVL
 EGAKQVPPQVQDILAMGGGGGPGYSQRRGMVNFDSGGGTQWDSGGGFGGGGG
 FSGGREGFGGREGGFGGREGGFGGRCMRDTSFGRRGNRGRGFTGPDAGHNVGG
 RGGGRCGRNNNMESRGFGGSGRGFGGRCMRDTSFGRRGNRGRGFTGPDAGHNVGG
 RSRSRGSYSRSRSRSRSRSRSRSRSRDRGGRGGRGGRGGRGGRGGRGGRGGRGGR
 VSGFDIKPPVSVVNLDMNAAAIAENNVPTSLSERQNGVGVSEEAALVPVVDPEP
 " 5843..5962
 /rpt_family="(TGG)n"
 complement(7077..>8090)
 /gene="F24P17.3"
 complement(7077..>8090)
 /gene="F24P17.3"
 /note="similar to hypothetical proteins GB:CAB16818,
 GB:AAD52277"
 complement(7254..8090)
 repeat_region
 mRNA
 complement(7077..>8090)
 gene
 /gene="F24P17.3"
 /note="similar to hypothetical proteins GB:CAB16818,
 GB:AAD52277"
 complement(7254..8090)
 CDS

ACCESSION U33007 271256
 VERSION U33007.1 GI:927685
 KEYWORDS
 SOURCE Saccharomyces cerevisiae strain-S288C (AB972).
 ORGANISM Saccharomyces cerevisiae
 Eukaryota: Fungi; Ascomycota: Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 REFERENCE 1 (bases 1 to 78500)
 AUTHORS Dietrich, F.S.
 TITLE The sequence of *S. cerevisiae* lambda 3641 and cosmids 9461, 9831, and 9410
 JOURNAL Unpublished (1995)
 REFERENCE 2 (bases 1 to 78500)
 AUTHORS Dietrich, F.S., Mulligan, J., Allen, E., Araujo, R., Aviles, E., Berrio, A., Carpenter, J., Chen, E., Cherry, J.M., Chung, E., Duncan, M., Hunnicke-Smith, S., Hyma, R., Komp, C., Lashkari, D., Lew, H., Lin, D., Mosedale, D., Nakahara, K., Namath, A., Oefner, P., Oh, C., Petel, F.X., Roberts, D., Schramm, S., Schroeder, M., Shogren, T., Shroff, N., Winant, A., Yelton, M., Botstein, D., and Davis, R.W.
 TITLE The sequence of *Saccharomyces cerevisiae* chromosome IV right
 JOURNAL Unpublished (1995)
 REFERENCE 3 (bases 1 to 78500)
 AUTHORS Dietrich, F.S.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-1995)
 REFERENCE 4 (bases 1 to 78500)
 AUTHORS Jia, Y. and Cherry, J.M.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUN-1997) Department of Genetics, Stanford University, Saccharomyces Genome Database, Stanford, CA 94305-5120, USA
 COMMENT
 Sequenced by: Stanford DNA Sequence & Technology Center
 855 California Avenue
 Palo Alto, CA 94304, USA
 Curated by: Saccharomyces Genome Database
 URL: <http://genome-www.stanford.edu/>
 e-mail: yeast-curator@genome.stanford.edu
 Neighboring Sequence:
 The 5' end of this sequence contains a 850 bp overlap with GenBank Accession Number U32274.
 Location/Qualifiers
 1..78500
 /organism="Saccharomyces cerevisiae"
 /strain="S288C (AB972)"
 /db_xref="taxon:4932"
 /chromosome="IV"
 /map="4P"
 /strain="S288C (AB972)"
 1..42652
 /organism="Saccharomyces cerevisiae"
 /db_xref="taxon:4932"
 /clone="cosmid 9461"
 /map="4P"
 /organism="Saccharomyces cerevisiae"
 /db_xref="taxon:4932"
 /clone="lambda 3641"
 complement(603..1322)
 /gene="STE14"
 complement(603..1322)
 /gene="STE14"
 /note="YDR410C; CAI: 0.12"
 /codon_start=1
 /product="Ste14p: farnesyl cysteine carboxyl-methyltransferase"
 /protein_id="AAB64880.1"
 /db_xref="GI:927714"
 /translation="MHODFOEDEHEYPDIRRNPLHEWTWTSVILGILLGIFVGLFPQI
 RFKNPLFIALLFHTLEYITAKYNPLKHSEFLNKGKSYNAHSAFLLLECLVE
 SFLFDPLKIFSYSLATKCTVLCGLVILGQYTRTAMTAGHSFHVTKKESDHV
 LVKTGVSWGRHPSYLGFFWMAIGTQLLLNPLSLVIFVFLWKFSDRIIRVEEKYLI
 EFFSAEYIEYKNKVGIGIPFI"
 complement(1600..2625)
 /gene="YDR411C"
 complement(1600..2625)
 /gene="YDR411C"
 /note="similar to *Caenorhabditis elegans* hypothetical protein from GenBank Accession Number U00036, CDS number 7; CAI: 0.10"
 /codon_start=1
 /evidence=not_experimental
 /product="Ydr411cp"
 /protein_id="AAB64889.1"
 /db_xref="GI:927723"
 /translation="MAGRNVRVRLHNGGRNNDVMGPKFELWLNIPTRTFLTLAIWM
 TIVGRNLINPWFYIYVNNITFKKQIWRLLTSCVMSLRAMPALMEIYSYDRSSOL
 ERGFGGLGNRRGPMVVDYAYLFCILATITATTIYGSYPVITISYFISCIY
 TSDIDNANQIMFYGLIPVGMGFLIQLFISGFNEGDFVLSLGTFTGLYLCLOT
 HTLGPIMGSMISRKADPTYGISPNKFTSPWFTSLYARITGAHNETAFFNNFANVS
 SORTRVFSGRGQRLGTAPATLSQTSITSDGRASGSQLRSFSLNQFQGRGQVGT
 NSPDSQ"
 2924..3631
 /gene="YDR412W"
 2924..3631
 /gene="YDR412W"
 /note="CAI: 0.16"
 /codon_start=1
 /evidence=not_experimental
 /product="Ydr412wp"
 /protein_id="AAB64855.1"
 /db_xref="GI:927689"
 /translation="MAVNTNRQILTRKNYATKQSKKFGTDEVTFDKSRDLYLTGFH
 KRKLQKQKAEFIKEQERLKIIEERQIRQERKEVMEQLTKFKESLNLEAIEDAK
 NDKTEDLQVDESDESHGFDSDKDDGNDNNESSVAPILKKGAITIYIDSTTVELET
 EPNDNFYLAQLNNVLEKAEKVLKQKINRAIKYAKFLGVDEKQKKPRVKRYLTK
 NERRINORKANDKKRRR"
 complement(3067..3642)
 /gene="YDR413C"
 complement(3067..3642)
 /gene="YDR413C"
 /codon_start=1
 /product="Ydr413cp"
 /protein_id="AAB64892.1"
 /db_xref="GI:2194165"
 /translation="MRLHYLLRLSPAFWLFILRSLVKYLNFFTLGFFGFCFSSTPRN
 FAYLVAFMLCLSTFSFSLNLTFSWAKYSKLSGNSVNSSTVSESSYISVIAPEFKI
 GFTLILSLSPSSLSSESQSDSCKSVLSFSPASSISASAKFLSLNWFNCSSI
 TSLRSCTIFCLSSITFLSLSCSLNSCAFFCL"
 complement(3829..4917)
 /gene="ERD1"
 complement(3829..4917)
 /gene="ERD1"
 /note="required from retention of luminal endoplasmic reticulum proteins; YDR414C; CAI: 0.11"
 /codon_start=1
 /product="Erd1p"
 /protein_id="AAB64888.1"
 /db_xref="GI:927722"
 /translation="MEKSESNSEGLYLQNLNVPQRFIVLIILALMTWILKFFL
 HSNLDVSQVILTRVPHDIRGYTLQQLHRTARNFALKITRIIPHFHFAIVLEFFMNI
 IEGPLKNILIVLEFLPIQCVTFWLLKCEQIIKYCTRCCLLIESPSRLNRYILI
 SDTLTSFGKPLIDFTLSLIFREPFTHFDLSVALLPLVLRLLQCLRYRLHEATLL
 FNALKYCNLPILFCTWRSRVSGSINEERLHVQWFMILINSYTLFWDVMDWSLD
 SLTSLRSRSKSAVTLKKKMYHSAILVDLFLRFLWLVYLSQNLKLVAAVDSYIFFOGE
 MOYEVIRGGIWWVVKLDAEYIKFASK"
 complement(5268..6392)
 /gene="YDR415C"
 complement(5268..6392)
 /gene="YDR415C"
 /note="similar to bacterial leucyl aminopeptidase from PIR Accession Number A38086; CAI: 0.14"
 /codon_start=1
 /evidence=not_experimental

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:26:37 ; Search time 22.15 Seconds
(without alignments)
1472.492 Million cell updates/sec

Title: US-09-271-584A-2

Perfect score: 2755
Sequence: 1 MLDSLVKPLSLSTSDHASV.....FVPFVPGSPTEPNPPDLKA 538

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_0601.*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2755	100.0	538	20	Arabidopsis thalia
2	2516	91.3	547	20	Arabidopsis thalia
3	2168.5	78.7	555	21	Atclplex gmelini N
4	2091.5	75.9	553	22	Protein regulating
5	2059.5	74.8	540	22	Protein regulating
6	2041	74.1	542	22	Protein regulating
7	2023.5	73.4	535	21	Rice Na+/H+ antipo
8	1982.5	72.0	555	22	Protein regulating
9	1570	57.0	529	20	Arabidopsis thalia
10	582	21.1	608	21	Cat flea HMT Na/H
11	565	20.5	526	22	Human secreted pro

12	564	20.5	633	20	AAV44017	Saccharomyces cere
13	563	20.4	509	22	AAV50555	Human secreted pro
14	544	19.7	569	20	AAV44016	Schizosaccharomyce
15	463	16.8	834	18	AAW11325	Human Na+/H+ excha
16	441	16.0	816	14	AAW44143	Rabbit sodium ion/
17	429	15.6	325	20	AAV40906	Arabidopsis thalia
18	320	11.6	339	22	AAV50590	Human secreted pro
19	320	11.6	339	22	AAV50590	Human secreted pro
20	263	9.5	232	20	AAV40907	Arabidopsis thalia
21	237	8.6	185	20	AAV40903	Arabidopsis thalia
22	151	5.5	79	21	AAV44993	Human secreted pro
23	144	5.2	511	22	AAV76843	Corynebacterium g1
24	143.5	5.2	178	20	AAV36089	Extended human sec
25	139	5.0	178	20	AAV35990	Extended human sec
26	137	5.0	480	19	AAW98431	H. pylori GRPO 446
27	136	4.9	129	21	AAW53670	Human colon cancer
28	129	4.7	412	21	AAV70156	Staphylococcus aur
29	119	4.3	443	21	AAV14094	Arabidopsis thalia
30	119	4.3	518	21	AAV14093	Arabidopsis thalia
31	119	4.3	568	21	AAV14092	Arabidopsis thalia
32	119	4.3	697	21	AAV93290	Amino acid sequenc
33	117.5	4.3	430	20	AAW88344	Salmonella enteric
34	117.5	4.3	469	18	AAW89797	Staphylococcus aur
35	117.5	4.3	482	13	AAV21409	NADH dehydrogenase
36	117	4.2	440	21	AAV13314	Arabidopsis thalia
37	117	4.2	440	21	AAV46728	Arabidopsis thalia
38	117	4.2	515	21	AAV13313	Arabidopsis thalia
39	117	4.2	515	21	AAV46727	Arabidopsis thalia
40	117	4.2	565	21	AAV13312	Arabidopsis thalia
41	117	4.2	565	21	AAV46726	Arabidopsis thalia
42	116	4.2	427	21	AAV46725	Arabidopsis thalia
43	116	4.2	502	21	AAV46724	Arabidopsis thalia
44	116	4.2	552	21	AAV46723	Arabidopsis thalia
45	113	4.1	407	19	AAV11034	H. pylori ORF 04ep

ALIGNMENTS

RESULT 1

AAV40901
ID AAY40901 standard; Protein: 538 AA.

XX AC AAY40901;

XX 18-JAN-2000 (first entry)

XX Arabidopsis thaliana Na/H transporter AtNHX1.

XX Sodium; proton; antiport; transporter; salt tolerance; salt management;
XX transgenic plant; survival; soil; farming; accumulation; irrigation;
XX crop.

OS Arabidopsis thaliana.

XX WO9947679-A2.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-CA00219.

XX 18-MAR-1998; 98US-0078474.

XX 15-JAN-1999; 99US-0116111.

XX (BLUM/) BLUMWALD E.

XX (APSE/) APSE M.

XX (SNED/) SNEDDEN W.

XX (AHAR/) AHARON G.

XX Blumwald E, Apse M, Snedden W, Aharon G;

XX WPI; 1999-571840/48.

XX N-PSDB; AA222591.

XX PT Nucleic acid molecules encoding sodium/proton transport polypeptides,
 XX PT useful in genetic engineering salt tolerance in crop plants -
 XX PS Claim 36; Fig 1A; 93pp; English.
 XX CC The invention relates to an isolated nucleic acid molecule encoding
 CC a plant Na/H antiporter (PNHX) transporter polypeptide, or a fragment
 CC and capable of increasing salt tolerance in a cell. This sequence
 CC corresponds to the AtNHX1 transporter from Arabidopsis thaliana.
 CC The Na/H transporter polypeptides provide a means of intracellular
 CC salt management, particularly in plants. The sequences are useful for
 CC producing transgenic plants that are capable of surviving in soil with
 CC high salt levels that would normally inhibit growth of the crop species.
 CC This would be useful in farming land in areas that are generally
 CC considered unproductive through salt accumulation and poor irrigation,
 CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
 CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
 CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
 CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
 CC increased salt tolerance.
 XX SQ Sequence 538 AA;

Query Match 100.0%; Score 2755; DB 20; Length 538;
 Best Local Similarity 100.0%; Pred. No. 3.9e-282;
 Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLDLSVSKLPSLSTDSHASVVALNLFVALLCACIVLGHLEENRMNSESITALLIGLGTG 60
 Db 1 mldlsvsklpslstshasvvalnlfvallcacivlgllleenrwmnesitalligltg 60
 QY 61 VTILLISKGKSHLLVFSDELFTYLLPPIIFNAGFOVKKKOFFRNFVTIMLFGAVGTII 120
 Db 61 vtilliskgkshllvfsedelffiyllppifnagfvkqkqffrnfvtimlfgavgtii 120
 QY 121 SCTIISLGVTQFFKKLDIGTFDGLDYLAIGAIFAAPDSVCTLOVNLQDETPLLVSFVGE 180
 Db 121 sctiislgvtqffkldigtfdglgdylaigaifaatdsctlvlnqdetppllvsfvg 180
 QY 181 GVVNDATSVVFNFAIQSFDTLHNEAAPHLLGNFLYFLSTLLGAATGLISAVYIKKL 240
 Db 181 gvvndatssvvfnfaigsfldthlnheaaphllgnflyflstllgaatglisavvikl 240
 QY 241 YFGRHSTDREVALMMLMAYLSYMLAEFLDLSGLTAVFFCGIWMSHYTHWNTVTESSRITTK 300
 Db 241 yfgrhstdrevalmmlmaylsymlaelfdlsiltvffcgimshytwntvtesrirttk 300
 QY 301 HTFATLSPLAETFTFLYVGMADLIDKWRVSPTPCTSIATSVSSILMGLVMVGRAAFVPL 360
 Db 301 htfatlsplaetftflvygmaldidkwrsvptpctsiatssilmglvmvgraafvpl 360
 QY 361 SFLSNLAKKNQSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTFRAGHTDVRGNALMITST 420
 Db 361 sflsnlakkngsekinfnmqvviwsglmrgavsmalaynknkfraghtdvrngalmitst 420
 QY 421 ITVCLFSFVVEGMLTKPLISYLLPHONATSMLSDDNTPKSTHPLDQDSFIEPSGNHN 480
 Db 421 itvclfsfvvegmktkplisyllphnqatmslssddntpksthipldqdsfiepsgnhn 480
 QY 481 VPRPDSIRGFLTRPTRVHYVYVQFDDSFMRPVFGGGRGVFPVPGSPTERNPPLDLSKA 538
 Db 481 vprpdsirgfltrptrvhyvyvqfddsfmrpvfgggrgvfpvpgsppternppdlska 538

RESULT 2

AA40905

ID AA40905 standard; Protein; 547 AA.

XX AC

XX AA40905;

XX DT

-18-JAN-2000 (first entry)

XX DE Arabidopsis thaliana Na/H transporter.
 XX KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
 KW transgenic plant; survival; soil; farming; accumulation; irrigation;
 KW crop.
 XX OS Arabidopsis thaliana.
 XX PN W09947679-A2.
 XX PD 23-SEP-1999.
 XX PF 18-MAR-1999; 99WO-CA00219.
 XX PR 18-MAR-1998; 98US-0078474.
 PR 15-JAN-1999; 99US-0116111.
 XX PA (BLUM/) BLUMWALD E.
 PA (APSE/) APSE M.
 PA (SNED/) SNEEDEN W.
 PA (AHAR/) AHARON G.
 XX PI Blumwald E, Apse M, Snedden W, Aharon G;
 XX WPI; 1999-571840/48.
 DR N-PSDB; AA222595.
 XX PT Nucleic acid molecules encoding sodium/proton transport polypeptides,
 PT useful in genetic engineering salt tolerance in crop plants -
 XX PS Claim 50; Fig 5A-B; 93pp; English.
 XX CC The invention relates to an isolated nucleic acid molecule encoding
 CC a plant Na/H antiporter (PNHX) transporter polypeptide, or a fragment
 CC and capable of increasing salt tolerance in a cell. This sequence
 CC corresponds to a transporter from Arabidopsis thaliana.
 CC The Na/H transporter polypeptides provide a means of intracellular
 CC salt management, particularly in plants. The sequences are useful for
 CC producing transgenic plants that are capable of surviving in soil with
 CC high salt levels that would normally inhibit growth of the crop species.
 CC This would be useful in farming land in areas that are generally
 CC considered unproductive through salt accumulation and poor irrigation,
 CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
 CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
 CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
 CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
 CC increased salt tolerance.
 XX SQ Sequence 547 AA;

Query Match 91.3%; Score 2516; DB 20; Length 547;
 Best Local Similarity 100.0%; Pred. No. 7e-257;
 Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLDLSVSKLPSLSTDSHASVVALNLFVALLCACIVLGHLEENRMNSESITALLIGLGTG 60
 Db 1 mldlsvsklpslstshasvvalnlfvallcacivlgllleenrwmnesitalligltg 60
 QY 61 VTILLISKGKSHLLVFSDELFTYLLPPIIFNAGFOVKKKOFFRNFVTIMLFGAVGTII 120
 Db 61 vtilliskgkshllvfsedelffiyllppifnagfvkqkqffrnfvtimlfgavgtii 120
 QY 121 SCTIISLGVTQFFKKLDIGTFDGLDYLAIGAIFAATDSVCTLOVNLQDETPLLVSFVGE 180
 Db 121 sctiislgvtqffkldigtfdglgdylaigaifaatdsctlvlnqdetppllvsfvg 180
 QY 181 GVVNDATSVVFNFAIQSFDTLHNEAAPHLLGNFLYFLSTLLGAATGLISAVYIKKL 240
 Db 181 gvvndatssvvfnfaigsfldthlnheaaphllgnflyflstllgaatglisavvikl 240
 QY 241 YFGRHSTDREVALMMLMAYLSYMLAEFLDLSGLTAVFFCGIWMSHYTHWNTVTESSRITTK 300

```

|||||
Db 241 yfgrhstdevalmmaylsymlaelfsgiltvfcgvmshytnvntessrittk 300
QY 301 HTFATLSFLAETFIPLYVGMADALDKWRSVSDTPTGTSIAVSSILMGLVMVGRAAFVPEPL 360
Db 301 htfatlsflaetfilyvgmdaldidkwrsvsdtpgtsiavsillmglvmvgraafvpl 360
QY 361 SFLSNLAKKNQSEKINFNMQVIVWWSGLMRGAVSMALAYNKFTTRAGHTDVRGNAMITST 420
Db 361 sflsnlakknsekinfnmqvivwsglmrgavsmalaynkfttraghtdvrgnalmist 420
QY 421 ITVCLFSTVVFVGMGLTKPLISYLLPHQNATTSMLSDDNTPKSIHIPLLD--QDSFIE-P 475
Db 421 itvclfstvfvfgmltkplisyllphqnattsmldsdnptksihiplld--qdsfiepsgnhn 480
QY 481 VPRPDSIRGFLRPRTR 496
Db 481 vprpdsirgflrprtr 496

RESULT 3
AAB12786
ID AAB12786 standard; Protein; 555 AA.
XX
AC AAB12786;
XX
DT 23-NOV-2000 (first entry)
XX
DE At1p1ex gmelini Na+ and H+ antiporter protein.
XX
KW At1p1ex gmelini; Na plus and H plus antiporter protein;
KW Na+ and H+ antiporter protein; transformed plant; high salt tolerance.
XX
OS At1p1ex gmelini.
XX
PN JP2000157287-A.
XX
PD 13-JUN-2000.
XX
PF 16-SEP-1999; 99JP-0261606.
XX
PR 24-SEP-1998; 98JP-0269504.
XX
PA (SHOK-) SHOKUBUTSU KOGAKU KK.
XX
DR WPI; 2000-468209/41.
DR N-PSDB; AAA72926.
XX
PT An Na+ and H+ plus antiporter protein and a gene encoding it
XX
PS Disclosure; Page 10-12; 16pp; Japanese.
XX
CC The present sequence represents an Na+ and H+ antiporter protein
CC isolated from At1p1ex gmelini. The Na+ and H+ antiporter protein and
CC gene encoding it are useful for the preparation of transformed plants
CC with high salt tolerance, e.g. for growth in arid land.
XX
SQ Sequence 555 AA;

Query Match 78.7%; Score 2168.5; DB 21; Length 555;
Best Local Similarity 77.7%; Pred. No. 3.7e-220;
Matches 425; Conservative 53; Mismatches 56; Indels 13; Gaps 5;

QY 2 LDSIVS-KLPSSLTSDHASVWALNLEVALLCACIVLGLHLEENRWMSITALLIGLGTG 60
Db 5 lssllsgkmdalttsdhasvsmnmfvalicgicvighleerwnmesitalliglatg 64
QY 61 VTILLISGKSHLLVFSEDLFFIYLLPPIIFNAGFQVKKQFFRNFTIMLFGAVGTII 120
Db 65 vvilllsgkshllvfseddlffiyllppiiifnagfqvkkqffnftivlfgavgtlv 124
QY 121 SCTIISLGVTQFPFKKLDIGTFDLGDYLAIGAIFAAFTSDVCTQLQVNLQDETPLLVSLEVE 180

```

```

|||||
Db 125 sftlslgalsifkklidgtleadylaigaafatdsvctqlvngdetpillsylvfe 184
QY 181 GVVNDATSVVFNATQSFDTLTHNHEAAFLHNLGNFLYLLSTLIGATGLISAYVKKL 240
Db 185 gvvndatsvvlfnatqsfldtridhrialqfmgfnflyfiastilgaftglisaylikkl 244
QY 241 YGRHSTDEVALMMLMAYLSYMLAELPDLSGILTVPFCGIVMSHYTHWNTESRITTK 300
Db 245 yfgrhstdevalmmaylsymlaelfylsgiltvfcgvmshytnvntessrvttk 304
QY 301 HTFATLSFLAETFIPLYVGMADALDKWRSVSDTPTGTSIAVSSILMGLVMVGRAAFVPEPL 360
Db 305 hafatlsfvaefilyvgmdaldiekwrfsdpsgiasvavssillglvmvgraafvpl 364
QY 361 SFLSNLAKKNQSEKINFNMQVIVWWSGLMRGAVSMALAYNKFTTRAGHTDVRGNAMITST 420
Db 365 swlmnfakksqekvtlqgvlvwwaglmrgavsmalaynqftrsghtqlrignalmist 424
QY 421 ITVCLFSTVVFVGMGLTKPLISYLLPHQNATTSMLSDDNTPKSIHIPLLD--QDSFIE-P 475
Db 425 isvlfstvmvfglktkplimfilpqpkhftscstvsdvgspsksyslpllegnqgdevdv 484
QY 476 SGNH-----NVPRPDSIRGFLRPRTRFVHYWRQFDDSEMRPVFGGRGVFPVPSPT 528
Db 485 ngnhedttrptivrpssirmalnaphthvhwkfkddsmrpvfggrgvfpvpspt 544
QY 529 ERNPPDL 535
Db 545 egstnl 551

RESULT 4
AAB73253
ID AAB73253 standard; Protein; 553 AA.
XX
AC AAB73253;
XX
DT 14-MAY-2001 (first entry)
XX
DE Protein regulating the pH of vacuoles.
XX
KW Vacuole pH regulation; flower colour.
XX
OS Nierembergia hybrida.
XX
PN WO200114560-A1.
XX
PD 01-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-JP05722.
XX
PR 24-AUG-1999; 99JP-0236800.
XX
PA (SUNR ) SUNTORY LTD.
XX
PI Iida S, Tanaka S, Inagaki Y;
XX
DR WPI; 2001-191648/19.
DR N-PSDB; AAF75765.
XX
PT Morning glory-originated gene encoding a protein with pH regulation
PT activity in vacuoles, useful in controlling flower color to give new
PT breeds of colorful plants for cut flowers, particularly applicable in
PT horticulture
XX
PS Example 7; Page 50-52; 68pp; Japanese.
XX
CC The present sequence is a protein, which has vacuolar pH regulatory
CC activities. The protein enables flower colour to be controlled via
CC regulation of the vacuolar pH, colours can range from blue to red in
CC colour spectrum. The protein is useful in controlling flower colour to
CC give new breeds of colourful plants for cut flowers, particularly

```

CC applicable in horticulture.

XX Sequence 553 AA;

Query Match 75.9%; Score 2091.5; DB 22; Length 553;

Best Local Similarity 76.3%; Pred. No. 4.9e-212; Mismatches 72; Indels 1; Gaps 1;

1;

QY 4 SLVSKLPSTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTGVTI 63

DB 7 tllgkmnltsdhsqsvsnlfvalicacivghlleenrwmmesitalvigstgvi 66

QY 64 LLSGKSHLLVFSEDFFIYLLPPIIFNAGFQVKKQFFNFVTIMLFGAVGTISCT 123

DB 67 lllsgkshllvfsedfiffyllpplifnagfvkksffrnfstlmfavggtllsfi 126

QY 124 IISLGVTFPKKLDIGFDGLYLAIGAFAATDSVCTLOVLNQDETPLLXSLVFGGVV 183

DB 127 lllsagaigifkmdighleigyilaigaifaatdsctqlvlnqeetppllyslvfegvv 186

QY 184 NDATSVVFNATOSFDLTHLNHEAFLHGNFLYLLPSTLLGAATGLISAVVKKLYFG 243

DB 187 ndatsvllfnavqnfldshlstgkalqlignflylfasstflgvavgllsafiikklyfg 246

QY 244 RHSTDREVALMMLVAYLSYMLAEFLDLSGILTVFFCGIVMSHYTWHNVTESSRITTKHTF 303

DB 247 rhstdrevalmmlaylsymlaelfylsgilvtvfcgivmshytwhnvteessrvttkhtf 306

QY 304 ATLSFLAETFTFLYVGMDALDIDKWRSDTPTGTSIAVSSILMGLVMVGRAAFVPLSFL 363

DB 307 atlsflaeftflfyvgmdaldiekwfkfvsdpptsikvssillglvlgvrgafvplsl 366

QY 364 SNLAKNQSEKINFNMQVVIWWSGLMRGAVSNALAYNKETFRAGHTDVRGNAMITSTTV 423

DB 367 snltkknpedktsfndqvtiwwaglmrgavsmalaynqftrgghqqlranaimitsttv 426

QY 424 CLFSTVVFGLTKPLISYLLPHQNAATSMLSDD-NTPKSIHIPLLDQDSEFIPSGNNHNP 482

DB 427 vlfstvvfglmtkpllllpskhlrmissepmtpksfivplldstqdsadlgrhvp 486

QY 483 RPSIRGFLTRPRTRVHYVWRQDDSFMRPVFGGRGFVPVPGSPTRNPP 533

DB 487 rphsirmllstpshtvhyvwrkfdnafmrpvfgrgfvpgpsptepvp 537

RESULT 5

AAB73252 standard; Protein; 540 AA.

XX

AC

XX

DT 14-MAY-2001 (first entry)

XX

DE Protein regulating the pH of vacuoles.

XX

KW Vacuole pH regulation; flower colour.

XX

OS Petunia hybrida.

XX

PN W0200114560-A1.

XX

PD 01-MAR-2001.

XX

PF 24-AUG-2000; 2000WO-JP05722.

XX

PR 24-AUG-1999; 99JP-0236800.

XX

PA (SUNR) SUNTORY LTD.

XX

PI Iida S, Tanaka S, Inagaki Y;

XX

DR WPI; 2001-191648/19.

N-PSDB; AAF75764.

XX Morning glory-originated gene encoding a protein with pH regulation activity in vacuoles, useful in controlling flower color to give new breeds of colorful plants for cut flowers, particularly applicable in horticulture -

XX Example 7; Page 42-45; 68pp; Japanese.

XX The present sequence is a protein, which has vacuolar pH regulatory activities. The protein enables flower colour to be controlled via regulation of the vacuolar pH, colours can range from blue to red in colour spectrum. The protein is useful in controlling flower colour to give new breeds of colourful plants for cut flowers, particularly applicable in horticulture.

XX Sequence 540 AA;

QY Query Match 74.8%; Score 2059.5; DB 22; Length 540;

Best Local Similarity 75.3%; Pred. No. 1.1e-208;

Matches 397; Conservative 57; Mismatches 72; Indels 1; Gaps 1;

QY 4 SLVSKLPSTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTGVTI 63

DB 7 tllgndvrlstsdhsqsvsnlfvalicacivghlleenrwmmesitalvigstgvi 66

QY 64 LLSGKSHLLVFSEDFFIYLLPPIIFNAGFQVKKQFFNFVTIMLFGAVGTISCT 123

DB 67 lllsgkshllvfsedfiffyllpplifnagfvkksffrnfstlmfalgltllsfi 126

QY 124 IISLGVTFPKKLDIGFDGLYLAIGAFAATDSVCTLOVLNQDETPLLXSLVFGGVV 183

DB 127 lllsagaigifkkmagileigyilaigaifaatdsctqlvlnqdetppllyslvfegvv 186

QY 184 NDATSVVFNATOSFDLTHLNHEAFLHGNFLYLLPSTLLGAATGLISAVVKKLYFG 243

DB 187 ndatsvvlfnagvqnfldshlstgkalqlignflylfasstflgvavgllsafiikklyfg 246

QY 244 RHSTDREVALMMLVAYLSYMLAEFLDLSGILTVFFCGIVMSHYTWHNVTESSRITTKHTF 303

DB 247 rhstdrevalmmlaylsymlaelfylsgilvtvfcgivmshytwhnvteessrvttkhtf 306

QY 304 ATLSFLAETFTFLYVGMDALDIDKWRSDTPTGTSIAVSSILMGLVMVGRAAFVPLSFL 363

DB 307 atlsflaeftflfyvgmdaldiekwfkfvsdpptsikvssillglvlgvrgafvplsl 366

QY 364 SNLAKNQSEKINFNMQVVIWWSGLMRGAVSNALAYNKETFRAGHTDVRGNAMITSTTV 423

DB 367 snltkktpeaktsfndqvtiwwaglmrgavsmalaynqftrgghqqlranaimitsttv 426

QY 424 CLFSTVVFGLTKPLISYLLP-HQNAATSMLSDDNTPKSIHIPLLDQDSEFIPSGNNHNP 482

DB 427 vlfstvvfglmtkplirllpskhlsmissepttpksfivplldstqdsadlerhvp 486

QY 483 RPSIRGFLTRPRTRVHYVWRQDDSFMRPVFGGRGFVPVPGSPTE 529

DB 487 rphsirmllstpshtvhyvwrkfdnafmrpvfgrgfvpgpsptd 533

RESULT 6

AAB73251

XX ID AAB73251 standard; Protein; 542 AA.

XX

AC AAB73251;

XX

DT 14-MAY-2001 (first entry)

XX

DE Protein regulating the pH of vacuoles.

XX

KW Vacuole pH regulation; morning glory; flower colour.

XX

OS Ipomoea nil.

XX WO200114560-A1.
PN 01-MAR-2001.
PD 24-AUG-2000; 2000WO-JP05722.
XX 24-AUG-1999; 99JP-0236800.
XX (SUNR) SUNTORY LTD.
XX Iida S, Tanaka S, Inagaki Y;
PI WPI; 2001-191648/19.
XX N-PSDB; AAF75752.
DR Morning glory-originated gene encoding a protein with pH regulation
DR activity in vacuoles, useful in controlling flower color to give new
XX breeds of colorful plants for cut flowers, particularly applicable in
XX horticulture -
XX
PS Claim 2; Page 30-33; 68pp; Japanese.
XX
XX The present sequence is a protein from Morning Glory, which has vacuolar
CC pH regulatory activities. The protein enables flower colour to be
CC controlled via regulation of the vacuolar pH, colours can range from blue
CC to red in colour spectrum. The protein is useful in controlling flower
CC colour to give new breeds of colourful plants for cut flowers,
CC particularly applicable in horticulture.
XX
XX SQ Sequence 542 AA;

Query Match 74.1%; Score 2041; DB 22; Length 542;
Best Local Similarity 73.8%; Pred. No. 1e-206;
Matches 399; Conservative 57; Mismatches 71; Indels 14; Gaps 4;
QY 2 LOSLVSKLPSSLSTSHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 61
DB 5 lssallqn-sdlftsdhasvsmnlfvallicacivlgllleenrwnesitalligltgv 63
QY 62 TILLISKGSSHLVFESEDLFFIYLLPPIIFNAGFOVKKQFFRNFVTIMLFGAVGTIIS 121
DB 64 vlllsggkshllvfesdlffiyllppilfnagfvkqkffvfmftlmfalgaltis 123
QY 122 CTIISLGVTQFFKKLDIGTFDGLDYLAIGAIFAATDSVCTIQVLNQDETPLYSLVFGE 181
DB 124 csisfgavkikfhldidfldfgylaigaifaatdsvctiqvlsqdetplyslvifge 183
QY 182 VVNDATSVVFNIAIOSFDLTHLNHEAAPHLLGNFLYLFLSTLLGAATGLISAYVIRKLY 241
DB 184 vvndatsvvlfnaiqsfdmstfdpklghfignflyflsftlfgvgiglicaylikly 243
QY 242 FGRHSTDRVALMMLMAYLSYMLAELFDLSGLTVPFCGIVMSHYTHWNVTSSRITTKH 301
DB 244 fgrhstdrevalmmlmsylsymlaelfylsgiltvfcgimshytwbnvtessrvtrh 303
QY 302 TFATLSFLAETFIPLYVGMDALDIDKWSVSDTPTCTSIATSVSILMGLVMVGRAAFVPLS 361
DB 304 sfatlsfvaetfilyvgmdaldiekwfvknsqglsvavssilvglilvgraaafvpls 363
QY 362 FLSNLAKNQSEKINFNMQVITWNSGLMRGAVSMALAYNKFTFRAGTDRVRGNAMITSTI 421
DB 364 flsnlakknsdkisfrqiiilwaglmrgavslalaynkfttsghtshenaimitstv 423
QY 422 TVCLSTVVFGMLTPTLSYLLP-----HQNATTSMKSDNDNTPKSIHIPLDQ--DS 471
DB 424 tvclstvvfvmktpkplnllpphkpmpsgshsmnt---sepspkhftvplldngpds 480
QY 472 FIEPSGNHNVPRPDSIRGFLTRPTRTVHYVWQFDSPMRPVFGGRGVFPVPGSPERN 531
DB 481 esdmktgpeavptalmrltrptthvrywrkfdsfmrpvfgrgfvfvgvsgpveqs 540

QY 532 P 532
DB 541 p 541
RESULT 7
AAB11832
ID AAB11832 standard; Protein; 535 AA.
XX AC AAB11832;
XX 07-NOV-2000 (first entry)
XX DE Rice Na+/H+ antiporter, OsnHX1.
XX OSNHX1; Na+/H+ antiporter; sodium/proton antiporter; countertransporter;
XX active transport; rice; transgenic plant; salt-tolerance.
XX Oryza sativa.
XX WO200037644-A1.
XX 29-JUN-2000.
XX 22-DEC-1999; 99WO-JP07224.
XX 22-DEC-1998; 98JP-0365604.
XX (NORQ) JAPAN MIN AGRIC FORESTRY & FISHERIES.
XX Fukuda A, Tanaka Y;
XX WPI; 2000-442672/38.
XX N-PSDB; AAA61876.
XX DNA encoding a sodium ion and proton counter-transporter protein of
PT rice origin for production of salt tolerant rice transformants -
XX
XX Claim 1a; Fig 1; 43pp; Japanese.
XX This sequence represents a novel rice Na+/H+ antiporter
CC (countertransporter), OsnHX1. The invention relates to OsnHX1 and
CC nucleic acids which encode it; vectors, host cells and transgenic plants
CC containing OsnHX1 nucleic acids; recombinant expression of OsnHX1; and
CC antibodies which recognise OsnHX1. OsnHX1 nucleic acids are useful in the
XX production of salt tolerant transgenic plants.
XX Sequence 535 AA;

Query Match 73.4%; Score 2023.5; DB 21; Length 535;
Best Local Similarity 73.3%; Pred. No. 7.2e-205;
Matches 389; Conservative 60; Mismatches 79; Indels 3; Gaps 2;
QY 2 LDSLVSKLPSSL-STDSHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 60
DB 3 mevaaralgaltysdyasvsnlfvallicacivlgllleenrwnesitalligltcg 62
QY 61 VTILLISKGSSHLVFESEDLFFIYLLPPIIFNAGFOVKKQFFRNFVTIMLFGAVGTII 120
DB 63 vvilmlmckgkshlrvfseidlffiyllppilfnagfvkqkqkfrnfmftlfgavgtml 122
QY 121 SCTIISLGVTQFFKKLDIGTFDGLDYLAIGAIFAATDSVCTIQVLNQDETPLYSLVFGE 180
DB 123 sfftislaaiaifsrmmigtldvgdflaigaifsatdsvctqlvlnqdetplyslvifge 182
QY 181 GVNDATSVVFNIAIOSFDLTHLNHEAAPHLLGNFLYLFLSTLLGAATGLISAYVIRKLY 240
DB 183 gvndatsvvlfnalnqfnldvdaavvklfnlgnffylfslsftlfgvfgaglsaylikl 242
QY 241 YFGRHSTDRVALMMLMAYLSYMLAELFDLSGLTVPFCGIVMSHYTHWNVTSSRITTK 300
DB 243 yfgrhstdrevalmmlmaylsymlaellldsgiltvfcgimshytwbnvtessrvtkk 302

OS XX PN XX PD XX PF XX PR XX PA XX PI XX DR DR

Claim 10: Page 891-893; 964pp; English.

The invention relates to novel cat flea (*Ctenocephalides felis*) nucleic acids which are expressed in hindgut and Malpighian tubule (HMT) tissue of cat flea, and nerve cord (HNC) tissue. The invention also relates to the encoded proteins. The invention additionally encompasses expression constructs, recombinant viruses and recombinant cells comprising the nucleic acids of the invention, recombinant production of the proteins, antibodies against the proteins, a method of identifying inhibitors of the proteins, and compositions comprising the inhibitors for the administration to an animal. The nucleic acids, and the proteins they encode may be used in the prevention, treatment and diagnosis of diseases associated with flea infestations. For example, the nucleic acids may be used to produce an HMT or HNC protein according to standard recombinant DNA methodology by inserting the nucleic acids into a host cell and culturing the cell to express the protein. The HMT and HNC nucleic acids may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect and quantitate the presence of cat flea or other homologous nucleic acid sequences in samples. They may also be used to study the expression and function of the proteins and their role in metabolism. The HMT and HNC proteins may be used as antigens in the production of specific antibodies, and in assays to identify modulators (agonists and antagonists) of HMT and/or HNC protein expression and activity. The anti-HMT/HNC protein antibodies and antagonists may also be used to downregulate protein expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of flea polypeptides in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The present sequence represents a cat flea HMT protein of the invention.

```

69 lpekgaeeqhsmsffvclvhalglllhlmlqtgfygylpessvvvfgallglln 128
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 LISKXSSHLL---VFSEDLGFYLLPPIFNAGFQVKKKOFFRNVITMLFAGVGTIIS 121
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129 lmskianwknecafstfvlvlpfifesgynlkhgfnqigsilvfaifgtai 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 CTIISLGVTFQFKKIDIGTFDILG-----DYLAIGALFAATDVCVLOVLNO-DETPLLY 174
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
189 afvvgagvyl-----lgnadvaynlsvfesafgslisadvpvaivfahldvdpvln 242
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
175 SLVFGCVNDATSVVVFNAI-QSFDLTHLNHEAAPHLLGNFLYLPFLSTLLGAAATGIS 233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
243 mlyfgesilndaislvtlaviesnnblmttaeeavvsgrlncrlmfafsaqlvgvfallis 302
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Cat flea; hindgut and Malpighian tubule protein; HMT;

Cat flea; hindgut and Malpighian tubule protein; HMT; flea infestation; vaccine; antiparasitic; therapeutic diagnosis; detection.

Cat flea; hindgut and Malpighian tubule protein; HMT; flea infestation; vaccine; antiparasitic; therapeutic diagnosis; detection.

OS Saccharomyces cerevisiae.
 PN WO9947679-A2.
 XX 23-SEP-1999.
 PD 18-MAR-1999; 99WO-CA00219.
 PF 18-MAR-1998; 98US-0078474.
 PR 15-JAN-1999; 99US-0116111.
 XX (BLUM/) BLOWMELD E.
 PA (APSE/) APSE M.
 PA (SNED/) SNEDDEN W.
 PA (AHAR/) AHARON G.
 XX Blumwald E, Apse M, Snedden W, Aharon G;
 XX WPI; 1999-571840/48.
 OR
 XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
 PT useful in genetic engineering salt tolerance in crop plants
 XX
 PS Disclosure; Fig 8B; 93pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule encoding
 CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
 CC and capable of increasing salt tolerance in a cell. This sequence
 CC corresponds to a Na/H antiport transporter from the budding yeast
 CC Saccharomyces cerevisiae. The Na/H transporter polypeptides provide
 CC a means of intracellular salt management, particularly in plants. The
 CC sequences are useful for producing transgenic plants that are capable
 CC of surviving in soil with high salt levels that would normally inhibit
 CC growth of the crop species. This would be useful in farming land in areas
 CC that are generally considered unproductive through salt accumulation and
 CC poor irrigation, e.g. in India, Australia, and prairies in USA or Canada.
 CC Commercial crops, such as potatoes, tomatoes, brassica, cotton,
 CC sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
 CC rye, barley, triplex, sorghum, alfalfa, salicornia and others would
 CC benefit from increased salt tolerance.
 XX
 XX Sequence 633 AA;
 SQ
 Query Match 20.5%; Score 564; DB 20; Length 633;
 Best Local Similarity 26.8%; Pred. NO. 1.4e-50;
 Matches 164; Conservative 121; Mismatches 200; Indels 128; Gaps 21;
 QY 14 TSDHASVVALNLFVALLCACIVLGHLEEN--RWNNESITALLIGLGTGVILLISKGKS 71
 Db 55 teemfswalfmllilalwssyyitqtkravrhetvisifgmviglirmspghyi 114
 QY 72 SHLVFSEDLFFIYLLPPIIFNAGFOVKKQKQFRNFVIMLFAGVGTIISCTIISLGVTO 131
 Db 115 qdvtfnussyfnnvllpplilnsgyelnqvnfnmllsillfaipgttisavvigi-ily 173
 QY 132 FPKKLDIGFDL--GDYLAIGAIFRATDSVCTLOVLNODER-PLLYSLVPGEGVVDNATS 188
 Db 174 iwfflgiesidlsfadamsvgalatobptvilsfnaykvdpklytifgeslndais 233
 QY 189 VVVFNAIQSFDTLPHLNEHAFLHGLNLFLLSTLLGAATGLISAYIKRYLFGFRHSTD 248
 Db 234 lvmfetcqkthgqpattssvfegagiflmtfsvlligvlgilvalllkhthirry-pq 292
 QY 249 REVALMMLWALYSYMLAEFLDLISGLITVFECGIVMSHVTWNVTSSTITTKHFAFLSP 308
 Db 293 iesclillayesffnsngchmsgvisllfcgltkhyaynmrsrrsrtiklyifqlar 352
 QY 309 LAETFFLYVGMDALDIDKWSRSDTPGTS-----IAVSSTILMLVMVGRAAFVFP 359
 Db 353 lsenffiygllelf-----levelvykplliivaal---sicvarwcavfp 396
 QY 360 LS-----FLSNLAKNQS--EKINFNQVVIWWSGLMRGAVSWALAYN---- 400

Db 397 lsqfwnviyrvktirsmgitsgenisvdpelipynqumtftwagl-rgavgvallalgiqe 455
 QY 401 -KFTRAGHTDVRGNAMITSTITVCLFSTVVFGMLTKPLISYLLPHQNATSMLSDDNT- 458
 Db 456 ykft-----llatlvvvvltviifggtagmlevl----niktgciseads 499
 QY 459 -----PKSTHI-----PLLDQDSFIEPSGNHNVPRPDSIRGFLT 492
 Db 500 ddefdieapraanllngssiqtdlpgysdnnspsididqf-avssnknlpnnlstdtgn 558
 QY 493 -----RPRTV-----HYWROFDSFMRPVPFGGGRGFVFP 523
 Db 559 fgglnetentspnparssmdknlrdkltifnsdsgwfgnfdgqvfkpvfld-nvpspl 617
 QY 524 PGSPTERNPPDLS 536
 Db 618 qdsatq-spadsf 629
 RESULT 13
 AAB90555
 ID AAB90555 standard; Protein; 509 AA.
 XX
 AC AAB90555;
 XX
 DT 01-JUN-2001 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 93.
 XX
 KW Human; secreted protein; immunomodulatory; antisclerotic;
 KW dermatological; antiinflammatory; anti-HIV; cytostatic; cardiant;
 KW vascular; anti-angiogenic; ophthalmological; neuroprotectant;
 KW neotropic; anticonvulsant; anti-alzheimers; antiparkinsonian;
 KW antimicrobial; vulnary; vaccine; gene therapy; cancer;
 KW protein coordinate data; infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200121658-A1.
 XX
 PD 29-MAR-2001.
 XX
 PF 22-SEP-2000; 2000WO-US26013.
 XX
 PR 24-SEP-1999; 99US-0155709.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
 PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
 PI Young PE, Wei P, Florence KA;
 XX
 DR WPI; 2001-235311/24.
 DR N-PSDB; AAF97895.
 XX
 PT Nucleic acids encoding 32 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 XX
 PS Claim 11; Page 779-780; 890pp; English.
 XX
 CC The present sequence is one of 32 novel human secreted polypeptides. The
 CC nucleic acid molecules and polypeptides may be used in the prevention,
 CC diagnosis and treatment of diseases such as immune disorders (e.g.
 CC multiple sclerosis, systemic lupus erythematosus and human
 CC immuno-deficiency virus (HIV) infections), hyperproliferative disorders
 CC (e.g. cancers and Gaucher's disease), cardiovascular diseases
 CC (e.g. scimitar syndrome, Chaga's cardiomyopathy and coronary
 CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
 CC infectious diseases and/or for promoting wound healing, regeneration

CC and/or chemotaxis. The nucleic acid molecules may be used to produce the
 CC secreted polypeptides. They may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acid
 CC sequences in samples. The polypeptides may be used as antigens in the
 CC production of antibodies and in assays to identify modulators of
 CC their expression and activity.

XX Sequence 509 AA;

Query Match 20.4%; Score 563; DB 22; Length 509;
 Best Local Similarity 31.1%; Pred. No. 1.3e-50;
 Matches 162; Conservative 90; Mismatches 165; Indels 104; Gaps 18;

QY 21 VALNLFVALLCAGTIVLGHLEEN--RMNSESITALLIGLGTGV----- 61
 DB 19 vxlltfillltltltlwtlkrvrlfhetglamyglivgvlrygtpatgrdksls 78
 QY 62 -----TILL-----ISKSSHL-----LVFSEDLFFIYLLPPI 90
 DB 79 ctqdrastfltnvnsqkffeytlkgeispgkinsveqndmrkvtfdpevfnnallppi 138
 QY 91 IFNAGVQKKQFFRNFVTIMLFGAVGTIISCTIIS---LGVTFQFKKLDIGTFDLGD-- 145
 DB 139 ifhagysikkrhfrnlgslayafgtaxscflignlmvgvklmklmg-----qlsdkf 194
 QY 146 ---YLAIGATFAATDSVCTLOVLNQDETPL-LYSLVFGGVVNDATSVVFNAIQSFDL 200
 DB 195 ytxxlfagalistdptvlaifnelhadvlyalifgesvindavaixssivayqp 254
 QY 201 THLNHE-----AAPHLLNFYLLFLLSTLGAAYGLISAVYIKKLYFGRHSTD-REVALM 254
 DB 255 aglnthafdaaaffkvgifigfsgftngavtgvtxvtxk--ftkxhxfllletalf 312
 QY 255 MLMAYLSYMLAELFDLSILTVFCGVVMSHYTHWNVTESRITTKHTATLSFLAETFI 314
 DB 313 flmswstflaeacgfgvavlfctqahylynnlsvesrtrkqlfevhflaenfi 372
 QY 315 FLYVGMADLIDKWRVSVDPTGTSIAVSSILMG---LVMVGRAAFVPLSFLSNIAKNO 371
 DB 373 fsynglalftfk-----hvfspifilgafvaiflgraaahyplsfllnlgrrh- 421
 QY 372 SEKNFNQVVIWWSGLMRGAVSMALAYNKFTRAGHTDVRGNAIMTSTITVCLFSTVVF 431
 DB 422 --kigwnfghmmfsgl-rqemafalair-----dasyarqmmfttllivftvwi 472
 QY 432 GMLTKPLISYLLPHQNAATSNLSDNDPTPKSIHPLLDQDSF 472
 DB 473 ggggttplmswl-----nirvgvdpdxdp-----pxxdstaf 504

RESULT 14

AA44016
 ID AA44016 standard; Protein; 569 AA.

XX AC AA44016;

XX DT 18-JAN-2000 (first entry)

XX DE Schizosaccharomyces pombe Na/H antiport transporter.

XX KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
 KW transgenic plant; survival; soil; farming; accumulation; irrigation;
 KW crop.

XX OS Schizosaccharomyces pombe.

XX PN WO9947679-A2.

XX PD 23-SEP-1999.

XX PF 18-MAR-1999; 99WO-CA00219.

PR 18-MAR-1998; 98US-0078474.
 PR 15-JAN-1999; 99US-0116111.
 XX (BLUM/) BLUMWALD E.
 PA (APSE/) APSE M.
 PA (SNED/) SNEDDEN W.
 PA (AHAR/) AHARON G.
 XX Blumwald E, Apse M, Snedden W, Aharon G;
 XX WPI; 1999-571840/48.
 DR Nucleic acid molecules encoding sodium/proton transport polypeptides,
 XX useful in genetic engineering salt tolerance in crop plants -
 PT Disclosure; Fig 8A; 93pp; English.
 XX The invention relates to an isolated nucleic acid molecule encoding
 CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
 CC and capable of increasing salt tolerance in a cell. This sequence
 CC corresponds to an Na/H antiport transporter from the fission yeast
 CC Schizosaccharomyces pombe. The Na/H transporter polypeptides provide
 CC a means of intracellular salt management, particularly in plants. The
 CC sequences are useful for producing transgenic plants that are capable
 CC of surviving in soil with high salt levels that would normally inhibit
 CC growth of the crop species. This would be useful in farming land in areas
 CC that are generally considered unproductive through salt accumulation and
 CC poor irrigation, e.g. in India, Australia, and prairies in USA or Canada.
 CC Commercial crops, such as potatoes, tomatoes, brassica, cotton, wheat,
 CC sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
 CC rye, barley, atriplex, sorghum, alfalfa, salicornia and others would
 CC benefit from increased salt tolerance.
 XX SQ Sequence 569 AA;

Query Match 19.7%; Score 544; DB 20; Length 569;

Best Local Similarity 31.7%; Pred. No. 1.5e-48;
 Matches 138; Conservative 103; Mismatches 157; Indels 38; Gaps 15;

QY 22 ALNLFVALLCAGTIVLGHLEEN--RMNSESITALLIGLGTGVTTLLISKKGK--SSHLLVFS 78
 DB 36 alfillvlllgalltgyyqskkairahetvisvfgvmvvgllirvspglliqmvsfh 94
 QY 79 EDLFFIYLLPPIFNAGVQKKQFFRNFVTIMLFGAVGTIISCTIISLGVTFQ----- 132
 DB 95 styfinvllppilnsgyelhgsnffnigtlttfaqtffisa--vtigvlyifsf 152
 QY 133 FKLDIGTFDLGDYLAIGATFAATDSVCTLOVLNQDET-PLLYSLVFGGVVNDATSVVV 191
 DB 153 fenlsm-tf--vealsmgatlsatdptvlaifnsykvdkklytiifgesilndavaivm 209
 QY 192 FNAIQSPDLTHLNHEAFLHLLGNFLYLLSTLGAATGLISAYVKKLYFGRHSTDREV 251
 DB 210 fetlqqfgqkthlffitfsgigifitffisilgvisglitalllkyslrryps-ies 268
 QY 252 ALMMLMAYLSYMLAELFDLSILTVFCGVVMSHYTHWNVTESRITTKHTATLSFLAE 311
 DB 269 ciillmaytyffsngchmsgvslfcgtikhyaffmsykalstkylfrvrlaqise 328
 QY 312 TFFLYVGMADL-IDKWRVSVDPTGTSIAVSSILMGLVWVGRAAFVPLSFLSNIAK- 369
 DB 329 nffifiyigmslftqvdlvykpfifilattvavta-----srymnvfpnlslakfhrq 380
 QY 370 ---NOSEKINFNQVVIWWSGLMRGAVSMALAYNKFTRAGHTDVRGNAIMTSTITVCLF 426
 DB 381 rngnlidhipysyqmmflwagl-rgavgvla-----agfegenaqlrat-tlvvvvl 432
 QY 427 STVVFGLTKPLISYL 442
 DB 433 tliifggtarmleil 448

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 09:52:16 ; Search time 155.97 Seconds
(without alignments)
8768.159 Million cell updates/sec

Title: US-09-271-584a-1

Perfect score: 2178
Sequence: 1 cctctgtttcttctctg.....aaaaaaaaaaaaaaaaaaaaa 2178

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2178	100.0	2178	20	Arabidopsis thaliana
2	2143.4	98.4	2284	20	Arabidopsis thaliana
3	832.8	38.2	1668	21	At1plex gmelini N
4	821.2	37.3	2553	22	Gene regulating th
5	813	37.3	2423	22	Gene regulating th
6	774.8	35.6	2237	22	Gene regulating th
7	768.8	35.3	2330	21	cDNA encoding rice
8	754.4	34.6	2361	22	Gene regulating th
9	523.8	24.0	1788	20	Arabidopsis thaliana
10	252.6	11.6	330	20	Arabidopsis Na/H a
11	226.4	10.4	378	20	Rice Na/H antiport

12	141.8	6.5	1824	21	AAC95372	Cat flea HMT Na/H
c 13	141.8	6.5	1824	21	AAC95373	Cat flea HMT Na/H
14	141.8	6.5	2080	21	AAC95370	Cat flea HMT Na/H
c 15	141.8	6.5	2080	21	AAC95371	Cat flea HMT Na/H
16	118.2	5.4	936	22	AAF58252	Oligonucleotide D1
17	118.2	5.4	936	22	AAF58253	Oligonucleotide D1
18	118.2	5.4	936	22	AAF58257	Oligonucleotide D1
19	118.2	5.4	936	22	AAF58259	Oligonucleotide D2
20	118.2	5.4	936	22	AAF58262	Oligonucleotide D2
21	118.2	5.4	936	22	AAF58255	Oligonucleotide D1
c 22	116.4	5.3	936	22	AAF58252	Oligonucleotide D1
c 23	116.4	5.3	936	22	AAF58254	Oligonucleotide D1
c 24	116.4	5.3	936	22	AAF58257	Oligonucleotide D1
c 25	116.4	5.3	936	22	AAF58259	Oligonucleotide D2
c 26	116.4	5.3	936	22	AAF58262	Oligonucleotide D2
c 27	116.4	5.3	936	22	AAF58255	Oligonucleotide D1
28	91.2	4.2	1581	22	AAF97895	Human secreted pro
29	82.2	3.8	1683	20	AAZ22602	Arabidopsis thaliana
30	64.4	3.0	1354	22	AAF97931	Human secreted pro
31	64.4	3.0	1688	22	AAF97930	Human secreted pro
c 32	62.6	2.9	244	22	AAF58238	Oligonucleotide D1
33	60.2	2.8	714	20	AAZ22593	Arabidopsis thaliana
34	60.2	2.8	2122	20	AAZ22603	Arabidopsis thaliana
35	58.6	2.7	268	20	AAZ22607	Rice Na/H antiport
36	57.2	2.6	244	22	AAF58238	Oligonucleotide D1
37	55.8	2.6	1244	20	AAZ97674	Extended human sec
38	55.4	2.5	1347	20	AAZ97773	Extended human sec
39	53.8	2.5	1291	21	AAA53975	TRIP-Ap53 tumour s
40	53.8	2.5	2007	19	AAV40754	C. felis esterase,
c 41	53.8	2.5	2007	19	AAV40755	C. felis esterase,
42	52.2	2.4	5852	12	AAQ11710	Dictyostelium plas
43	52	2.4	263	21	AAA44056	Human secreted exp
44	48.8	2.2	682	21	AAA44930	Human secreted exp
45	48.8	2.2	1169	20	AAV99983	3' untranslated re

ALIGNMENTS

RESULT 1
AAZ22591
ID AAZ22591 standard; cDNA; 2178 BP.
XX
AC AAZ22591;
XX
DT 18-JAN-2000 (first entry)
XX
DE Arabidopsis thaliana Na/H transporter gene AtNHX1.
XX
KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
transgenic plant; survival; soil; farming; accumulation; irrigation;
crop: ss.
XX
OS Arabidopsis thaliana.
XX
PN WO9947679-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-CA00219.
XX
PR 18-MAR-1998; 98US-0078474.
PR 15-JAN-1999; 99US-0116111.
XX
PA (BLUM/) BLUMWALD E.
PA (APSE/) APSE M.
PA (SNED/) SNEDDEN W.
PA (AHAR/) AHARON G.
XX
PI Blumwald E, Apse M, Snedden W, Aharon G;
XX
DR WPI: 1999-571840/48.
DR P-PSDB; AAY40901.

XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
PT useful in genetic engineering salt tolerance in crop plants -
XX
PS Claim 4; Fig 1A; 93pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule encoding a
CC plant Na/H antiport (PNHX) transporter polypeptide, or a fragment and
CC capable of increasing salt tolerance in a cell. This sequence corresponds
CC to the gene encoding the AtNHX1 transporter from Arabidopsis thaliana.
CC The Na/H transporter polypeptides provide a means of intracellular salt
CC management, particularly in plants. The sequences are useful for
CC producing transgenic plants that are capable of surviving in soil with
CC high salt levels that would normally inhibit growth of the crop species.
CC This would be useful in farming land in areas that are generally
CC considered unproductive through salt accumulation and poor irrigation,
CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
CC increased salt tolerance.
XX
SQ Sequence 2178 BP; 576 A; 413 C; 472 G; 717 T; 0 other;

Query Match 100.0%; Score 2178; DB 20; Length 2178;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctctgtttctgtctctgtacgacgaagaagaagaatcctcaggttttagcttcga 60
DB 1 cctctgtttctgtctctgtacgacgaagaagaagaatcctcaggttttagcttcga 60

QY 61 agcttcacaaatttgaaatttgatctctctgtggtcttttgaatcagactaagatat 120
DB 61 agcttcacaaatttgaaatttgatctctctgtggtcttttgaatcagactaagatat 120

QY 121 ttgattaccacagaagtgttcaagaagaattgttcagtgagacacggaagaataaag 180
DB 121 ttgattaccacagaagtgttcaagaagaattgttcagtgagacacggaagaataaag 180

QY 181 agacttttttccagattttgctgatccaaatctgaaatagttgttcaattcttgat 240
DB 181 agacttttttccagattttgctgatccaaatctgaaatagttgttcaattcttgat 240

QY 241 caaatctgaaagagaagtgttggttgatctagaagaagatacaaatgttgattctcta 300
DB 241 caaatctgaaagagaagtgttggttgatctagaagaagatacaaatgttgattctcta 300

QY 301 gtgtcgaaactgccttcgttatcgacatctgacgcgttcgtgtggtgcgtgaatctc 360
DB 301 gtgtcgaaactgccttcgttatcgacatctgacgcgttcgtgtggtgcgtgaatctc 360

QY 361 ttgtgtgcaacttttgtgctgtattgtttgttcattcttttgaagagaatagatgg 420
DB 361 ttgtgtgcaacttttgtgctgtattgtttgttcattcttttgaagagaatagatgg 420

QY 421 atgaacgaatccatccacgccttctgtgattgggttaggcacgtggtgtaccattttgtg 480
DB 421 atgaacgaatccatccacgccttctgtgattgggttaggcacgtggtgtaccattttgtg 480

QY 481 attagttaaaagaaaagctcgacatctctcgtctttagtggaagatcttttctcatat 540
DB 481 attagttaaaagaaaagctcgacatctctcgtctttagtggaagatcttttctcatat 540

QY 541 cttttgccccattatttcaatcaggggtttcaagtataaaagaagcagtttttccgcg 600
DB 541 cttttgccccattatttcaatcaggggtttcaagtataaaagaagcagtttttccgcg 600

QY 601 aattctgtactattatgtcttttggctgtgtgggaactattatttcttgacaaatcata 660
DB 601 aattctgtactattatgtcttttggctgtgtgggaactattatttcttgacaaatcata 660

QY 661 tctctaggtgtaacacagattcttttaagaagttggacattggaaacctttgacttgggtgat 720
DB 661 tctctaggtgtaacacagattcttttaagaagttggacattggaaacctttgacttgggtgat 720

QY 721 tatcttgctattggtgcatatttctgtgcaacagattcagatgtacactgcaggttctg 780
DB 721 tatcttgctattggtgcatatttctgtgcaacagattcagatgtacactgcaggttctg 780

QY 781 aatcaacacagacacacatttcttaccagctctgtattcggagaggtgttctgaatgat 840
DB 781 aatcaacacagacacacatttcttaccagctctgtattcggagaggtgttctgaatgat 840

QY 841 gcaactcagttgtggtcttcaacgogattcagaggtttgatctcactcacctaacaac 900
DB 841 gcaactcagttgtggtcttcaacgogattcagaggtttgatctcactcacctaacaac 900

QY 901 gaactgcttttcatcttcttgaaactcttctatttcttctcctaagtaaccttgcct 960
DB 901 gaactgcttttcatcttcttgaaactcttctatttcttctcctaagtaaccttgcct 960

QY 961 ggtgtgcaacccggtctgataaagtgcgtatgttatcaagaagctatatttggaggcac 1020
DB 961 ggtgtgcaacccggtctgataaagtgcgtatgttatcaagaagctatatttggaggcac 1020

QY 1021 tcaactgaccgagaggttgccttatgatcttatgcttatgcttatcttcttatctgct 1080
DB 1021 tcaactgaccgagaggttgccttatgatcttatgcttatgcttatcttcttatctgct 1080

QY 1081 gactgtttctgacttgacggttatctcactcactgtttttctgtgtgtattgtgatgccat 1140
DB 1081 gactgtttctgacttgacggttatctcactcactgtttttctgtgtgtattgtgatgccat 1140

QY 1141 tacacatggcacaatgtaacgagagctcaagaatacaacaagaacataccttgcacat 1200
DB 1141 tacacatggcacaatgtaacgagagctcaagaatacaacaagaacataccttgcacat 1200

QY 1201 ttgtctattcttcggagacatttatcttctgtattgttgaaatggatgcttggacatt 1260
DB 1201 ttgtctattcttcggagacatttatcttctgtattgttgaaatggatgcttggacatt 1260

QY 1261 gacaagtggagatccgtgagtgcacacccgggaacatcgatcgagtgagctcaactcta 1320
DB 1261 gacaagtggagatccgtgagtgcacacccgggaacatcgatcgagtgagctcaactcta 1320

QY 1321 atgggtctgtctatggttggagagacgcttcttccgttatcgttctctatctaac 1380
DB 1321 atgggtctgtctatggttggagagacgcttcttccgttatcgttctctatctaac 1380

QY 1381 tttagccaagaagaaatcaaaagcgagaaaaatcaactttaacatgcaggttggattggtg 1440
DB 1381 tttagccaagaagaaatcaaaagcgagaaaaatcaactttaacatgcaggttggattggtg 1440

QY 1441 tctggtctcatgagaggtgctgtatctatgctcttgatatacaacaagaatttacaaggcc 1500
DB 1441 tctggtctcatgagaggtgctgtatctatgctcttgatatacaacaagaatttacaaggcc 1500

QY 1501 gggcacacagatgacgcggtggaatgcaatcatgatcaacgatacgaataactgtctgtctt 1560
DB 1501 gggcacacagatgacgcggtggaatgcaatcatgatcaacgatacgaataactgtctgtctt 1560

QY 1561 tttagcacagtggttgttggatgctgacccaaacacactcataagctactattaccgcac 1620
DB 1561 tttagcacagtggttgttggatgctgacccaaacacactcataagctactattaccgcac 1620

QY 1621 cagaaacgcccacacagacatgttatctgatgacacaccccaaaatccatcatatccct 1680
DB 1621 cagaaacgcccacacagacatgttatctgatgacacaccccaaaatccatcatatccct 1680

QY 1681 ttgttgaccacaaactcgttctattgagccttcagggaaccacaaatgtgctcgccgtgac 1740
DB 1681 ttgttgaccacaaactcgttctattgagccttcagggaaccacaaatgtgctcgccgtgac 1740

QY 1741 agtatcgtgtgtcttcttgacacgcccactcgaaacgctgcattactactggagacaaattt 1800

Db 1741 agtatacgtgctcttgacacgcccactcgaacgctgcattactacggacaaattt 1800
 QY 1801 gatgactccctcagcagaccgctcttgaggctgctggttgtaacctttgtccaggt 1860
 Db 1801 gatgactccctcagcagaccgctcttgaggctgctggttgtaacctttgtccaggt 1860
 QY 1861 tctccactgagagaacccctcctgatcttagtaagccttgagggttaacgtggaagaaa 1920
 Db 1861 tctccactgagagaacccctcctgatcttagtaagccttgagggttaacgtggaagaaa 1920
 QY 1921 gctttgatttttttgtagaagaagggtgattcaaatattgcttttggtaattatcca 1980
 Db 1921 gctttgatttttttgtagaagaagggtgattcaaatattgcttttggtaattatcca 1980
 QY 1981 ttgttaataattgtgtgagacagaaaatcgtctcctaaacgtttttgagagcagaagcaaa 2040
 Db 1981 ttgttaataattgtgtgagagcagaaaatcgtctcctaaacgtttttgagagcagaagcaaa 2040
 QY 2041 acatggcaactttgaagtggttgattgatgtatgtaattattattcatattttgtgtg 2100
 Db 2041 acatggcaactttgaagtggttgattgatgtatgtaattattattcatattttgtgtg 2100
 QY 2101 taacacaaactacacattgtttgtattgttttgattggtttttgttcgaaaaaaaaa 2160
 Db 2101 taacacaaactacacattgtttgtattgttttgattggtttttgttcgaaaaaaaaa 2160
 QY 2161 aaaaaaaaaaaaaaaa 2178
 Db 2161 aaaaaaaaaaaaaaaa 2178

RESULT 2

AAZ22595
 ID AAZ22595 standard; cDNA; 2284 BP.

XX AC AAZ22595;
 XX AC AAZ22595;
 DT 18-JAN-2000 (first entry)
 DE Arabidopsis thaliana Na/H transporter gene.
 KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
 KW transgenic plant; survival; soil; farming; accumulation; irrigation;
 KW crop; ss.
 OS Arabidopsis thaliana.
 XX PN W09947679-A2.
 XX PD 23-SEP-1999.
 XX PF 18-MAR-1999; 99WO-CA00219.
 XX PR 18-MAR-1998; 98US-0078474.
 XX PR 15-JAN-1999; 99US-0116111.
 XX PA (BLUM/) BLUMWALD E.
 XX PA (APSE/) APSE M.
 XX PA (SNED/) SNEDDEN W.
 XX PA (AHAR/) AHARON G.
 XX PI Blumwald E, Apse M, Snedden W, Aharon G;
 XX WPI; 1999-571840/48.
 XX DR P-PSDB; AAY40905.
 XX PT Nucleic acid molecules encoding sodium/proton transport polypeptides,
 XX useful in genetic engineering salt tolerance in crop plants -
 XX Claim 48; Fig 5A-B; 93pp; English.
 XX The invention relates to an isolated nucleic acid molecule encoding

CC a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
 CC and capable of increasing salt tolerance in a cell. This sequence
 CC corresponds to the gene encoding a transporter from Arabidopsis thaliana.
 CC The Na/H transporter polypeptides provide a means of intracellular salt
 CC management, particularly in plants. The sequences are useful for
 CC producing transgenic plants that are capable of surviving in soil with
 CC high salt levels that would normally inhibit growth of the crop species.
 CC This would be useful in farming land in areas that are generally
 CC considered unproductive through salt accumulation and poor irrigation,
 CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
 CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
 CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, barley,
 CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
 CC increased salt tolerance.

XX Sequence 2284 BP; 589 A; 444 C; 485 G; 766 T; 0 other;

Query Match 98.4%; Score 2143.4; DB 20; Length 2284;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2166; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 cctctgtttcgtctcgtagacgaagaagaagaatctcaggttttagcttcga 60
 Db 116 cctctgtttcgtctcgtagacgaagaagaagaatctcaggttttagcttcga 175
 QY 61 agcttccaaaaatttgattgtatctctcgtcctcttttgaatcagactgaagatat 120
 Db 176 agcttccaaaaatttgattgtatctctcgtcctcttttgaatcagactgaagatat 235
 QY 121 ttgattaccacagaagttgttcaaggaaatgtttcagtgacagcagcaggaataaaag 180
 Db 236 ttgattaccacagaagttgttcaaggaaatgtttcagtgacagcagcaggaataaaag 295
 QY 181 agacttttttccagattttgctgacccaaatctgaatagttgttcattgtcttgat 240
 Db 296 agacttttttccagattttgctgacccaaatctgaatagttgttcattgtcttgat 355
 QY 241 caaatctggaagagaagttgttgatctagaaagatacaaatgttggattctcta 300
 Db 356 caaatctggaagagaagttgttgatctagaaagatacaaatgttggattctcta 415
 QY 301 gtgtcgaaactgcttcgttattcgacactgacacgttctgtggttcgttgaaatctc 360
 Db 416 gtgtcgaaactgcttcgttattcgacactgacacgttctgtggttcgttgaaatctc 475
 QY 361 ttgtgtcactttttgtgtgtattgttcttggctcactcttttggagaagataagatgg 420
 Db 476 ttgtgtcactttttgtgtgtattgttcttggctcactcttttggagaagataagatgg 535
 QY 421 atgaacgaatccatccacgcttggattgggttagggaactggtgttaccatttggctg 480
 Db 536 atgaacgaatccatccacgcttggattgggttagggaactggtgttaccatttggctg 595
 QY 481 attagtaagggaaaaagctcgcatcttctcgtctttagtgaagatctttcttcacatat 540
 Db 596 attagtaagggaaaaagctcgcatcttctcgtctttagtgaagatctttcttcacatat 655
 QY 541 cttttgccaccattattatccaatgcagggtttccaagtaaaaaaagcagtttttccgc 600
 Db 656 cttttgccaccattattatccaatgcagggtttccaagtaaaaaaagcagtttttccgc 715
 QY 601 aatttcgtgactattatgctttttgtgctgttgggaactattattcttgcacaataca 660
 Db 716 aatttcgtgactattatgctttttgtgctgttgggaactattattcttgcacaataca 775
 QY 661 tctctagggtgaacacagttctttaaagaagttggacattggaaacctttgacttgggtgat 720
 Db 776 tctctagggtgaacacagttctttaaagaagttggacattggaaacctttgacttgggtgat 835
 QY 721 tatcttgcatttgggtgccattatttgcgcacagattcagatgtacacgtgcaggttctg 780
 Db 836 tatcttgcatttgggtgccattatttgcgcacagattcagatgtacacgtgcaggttctg 895

QY 781 aatcaagcagagacaccttttgcctttatacagttctgtattcgcgagaggtgtgtgtaaatgat 840
Db 896 aatcaagcagagacaccttttgcctttatacagttctgtattcgcgagaggtgtgtgtaaatgat 955
QY 841 gcaacgtcagttgtgtctcaacgcgattcagagctttgatctcaactcaactaaacac 900
Db 956 gcaacgtcagttgtgtctcaacgcgattcagagctttgatctcaactcaactaaacac 1015
QY 901 gaagctgtcttttcattctcttcttggaacctcttctgtattgtttctctcctaagtacaccttgctt 960
Db 1016 gaagctgtcttttcattctcttcttggaacctcttctgtattgtttctctcctaagtacaccttgctt 1075
QY 961 ggtgtcgaacccggtctctataagtcgcttatattcaagaagcctataacttttggaagccac 1020
Db 1076 ggtgtcgaacccggtctctataagtcgcttatattcaagaagcctataacttttggaagccac 1135
QY 1021 tcaactgaccgagaggttgcctcttatgatcttattggtcgtatcttcttcttatatgcttgcct 1080
Db 1136 tcaactgaccgagaggttgcctcttatgatcttattggtcgtatcttcttcttatatgcttgcct 1195
QY 1081 gagcttttgcacttgagcgggtatctctactgtgttttctctgtgtattgtgtatgtatcccat 1140
Db 1196 gagcttttgcacttgagcgggtatctctactgtgttttctctgtgtattgtgtatgtatcccat 1255
QY 1141 tacacatgccaatgttaacgagagctcaagaatacaacaagaacataccttttgcaact 1200
Db 1256 tacacatgccaatgttaacgagagctcaagaatacaacaagaacataccttttgcaact 1315
QY 1201 ttgtcatttcttcgagagacattttatttctgtatgttggaaaggatgccttggacatt 1260
Db 1316 ttgtcatttcttcgagagacattttatttctgtatgttggaaaggatgccttggacatt 1375
QY 1261 gacaagtggagatcogtgagtgacacacgggaacatcgatcgagtgagtcgaatcccta 1320
Db 1376 gacaagtggagatcogtgagtgacacacgggaacatcgatcgagtgagtcgaatcccta 1435
QY 1321 atgggtctgtcattggttggagagcagcgttctcttccgtttatcgtttcttatctaac 1380
Db 1436 atgggtctgtcattggttggagagcagcgttctcttccgtttatcgtttcttatctaac 1495
QY 1381 ttagccaagaagaatacaagcagagaaaaatcaactttaacatgcaggttgtgatttgggtgg 1440
Db 1496 ttagccaagaagaatacaagcagagaaaaatcaactttaacatgcaggttgtgatttgggtgg 1555
QY 1441 tctggtctcatgagaggtgctgtatctatgctcttgctcttgatatacaacaagtttacaagggcc 1500
Db 1556 tctggtctcatgagaggtgctgtatctatgctcttgctcttgatatacaacaagtttacaagggcc 1615
QY 1501 gggcacacagatgtacgcgggaatgcaatgcatgatacagagtaacgataactgtctgtctt 1560
Db 1616 gggcacacagatgtacgcgggaatgcaatgcatgatacagagtaacgataactgtctgtctt 1675
QY 1561 tttagacagtggtgttggatgtcgtacaaaacacactaagctacactattacgcgcac 1620
Db 1676 tttagacagtggtgttggatgtcgtacaaaacacactaagctacactattacgcgcac 1735
QY 1621 cagaacgcccaccagcagctgttatctgatgacaaccccccaaaatccatacatatccct 1680
Db 1736 cagaacgcccaccagcagctgttatctgatgacaaccccccaaaatccatacatatccct 1795
QY 1681 ttgttgaccaagaactcgttctcatgttagccttcagggaacacacaatgtgcctcggcctgac 1740
Db 1796 ttgttgaccaagaactcgttctcatgttagccttcagggaacacacaatgtgcctcggcctgac 1855
QY 1741 agtatcagtggtcttcttgacacggccactc-gaacctgtcattact-actggagacaaat 1798
Db 1856 agtatcagtggtcttcttgacacggccactcggaaacccggaacccgtaactaactaggagacaaat 1915
QY 1799 ttgatgactcttctcatgcgaccgcgtctttggaggtcgttggcttggactccttggcttccag 1858
Db 1916 ttgatgactcttctcatgcgaccgcgtctttggaggtcgttggcttggactccttggcttccag 1975

QY 1859 gttctccaactgagagaaacccctcctgatcttagtaagcgttgagggtgaacgtggaagaa 1918
Db 1976 gttctccaactgagagaaacccctcctgatcttagtaagcgttgagggtgaacgtggaagaa 2035
QY 1919 aagctttgattttttttgtgtagaaaaagggtgattcaaaattatgcttttgtgtaaaattatc 1978
Db 2036 aagctttgattttttttgtgtagaaaaagggtgattcaaaattatgcttttgtgtaaaattatc 2095
QY 1979 catttgtataattgtttgtgagacagaaaatctgtcttaacgtttttgagagcagaaagca 2038
Db 2096 catttgtataattgtttgtgagacagaaaatctgtcttaacgtttttgagagcagaaagca 2155
QY 2039 aaacatggcaacctttgaagtgtttgatgtatgtatgaattatattcatatttgtttgt 2098
Db 2156 aaacatggcaacctttgaagtgtttgatgtatgtatgaattatattcatatttgtttgt 2215
QY 2099 tgtaaacacaaactacacatttgtttatgttttgaattgttttgcctcgaaaaaaaaa 2158
Db 2216 tgtaaacacaaactacacatttgtttatgttttgaattgttttgcctcgaaaaaaaaa 2275
QY 2159 aaaaaaaaaa 2167
Db 2276 aaaaaaaaaa 2284
RESULT 3
AAAT72926
ID AAA72926 standard; DNA; 1668 BP.
XX AC AAA72926;
XX XX
DT 23-NOV-2000 (first entry)
XX DE Atliplex gmelini Na+ and H+ antiporter protein encoding DNA.
XX KW Atliplex gmelini; Na plus and H plus antiporter protein;
KW Na+ and H+ antiporter protein; transformed plant;
XX high salt tolerance; ds.
XX OS Atliplex gmelini.
XX PN JP2000157287-A.
XX PD 13-JUN-2000.
XX PF 16-SEP-1999; 99JP-0261606.
XX PR 24-SEP-1998; 98JP-0269504.
XX PA (SHOK-) SHOKUBUTSU KOGAKU KK.
XX WPI; 2000-468209/41.
DR P-PSDB; AAB12786.
XX PT An Na+ and H+ plus antiporter protein and a gene encoding it
XX PS Claim 1; Page 9-10; 16pp; Japanese.
XX CC The present sequence encodes an Na+ and H+ antiporter protein isolated
CC from Atliplex gmelini. The Na+ and H+ antiporter protein and gene
CC encoding it are useful for the preparation of transformed plants with
CC high salt tolerance, e.g. for growth in arid land.
XX XX
SQ Sequence 1668 BP; 397 A; 342 C; 376 G; 553 T; 0 other;

Query Match 38.2%; Score 832.8; DB 21; Length 1668;
Best Local Similarity 71.9%; Pred. No. 1.1e-176;
Matches 1146; Conservative 0; Mismatches 412; Indels 36; Gaps 3;
QY 323 cgacatctgatacagccttctgtgttcggttggaatctcttcttggacactcttcttgcctt 382
Db 50 ccaactctgatacagccttctgtgttcggttcgatacgaactgtttgtggaactgttatgtggtt 109

Qy	1463	tatctatggctcttgatacaacaagtttaacaaggcgccagacacagatgtatccgcggga	1522
Db	1490	ttctcatggcaacttgcttatactaatcagttttacagagcttgggcacacacagcttagggaa	1549
Qy	1523	atgcaatcatgatcacagatgacgataaactgtctgtcttttttagcacagtggtgtttggta	1582
Db	1520	atgcaatcatgatcaacaagcaactatctgtgtccttttttagtacaatgggtgtctcggt	1509
Qy	1583	tgctgacaaacacactcataaagctacctattaccgcaccagaacgcgcacac-----ga	1636
Db	1610	tgctgacaagcctctcatcatgtttttgtgctcctcaacgaaacacttcaactagtgtca	1669
Qy	1637	gcattgtatctgatgacaacaccccaaat-----cc	1668
Db	1670	gcacgcatcagatgtgggaagtcctaaagtcatactctgtgcacctcttgagggcacac	1729
Qy	1669	atacatatccctttgttgaccaagactcgttcattgagccttcagggaaacac--aatg	1726
Db	1430	aagattatgaagtgtatgtgtgggaacgcgaaaccatgagacaccacitgagcgcggacta	1489
Qy	1727	tgctcggcctgcacagtaatacgttgctcttttgacacggcccaactgcgaacctgacttact	1786
Db	1490	tagttgcacctagtagcctccgatcgtctctataatgacacctactcaacacgctccatcact	1549
Qy	1787	actgagacaaatttgatgactccttcacgcaccgcgtttttggaggctggtgctttgtac	1846
Db	1550	atfggcgcaaatcgatgactcctcatcgccgcgctttttgtgtggccgggtttttgtac	1609
Qy	1847	ccttgttccaggattctccaactgagagaaaccc	1880
Db	1610	ctttgtccccggattctcacctactcaacaaagac	1643

RESULT 4

AAAF75765	standard; DNA; 2553 BP.
AAAF75765;	
14-MAY-2001	(first entry)
Gene regulating the pH of vacuoles.	
Vacuole pH regulation; flower colour; ds.	
Nierembergia hybrida.	
WO200114560-A1.	
01-MAR-2001.	
24-AUG-2000; 2000WO-JP05722.	
24-AUG-1999; 99JP-0236800.	
(SUNR) SUNTORY LTD.	
Iida S, Tanaka S, Inagaki Y;	
WPI; 2001-191648/19.	
P-PSDB; AAB73253.	
Morning glory-originated gene encoding a protein with pH regulation activity in vacuoles, useful in controlling flower color to give new breeds of colorful plants for cut flowers, particularly applicable in horticulture -	
Example 7; Page 45-49; 68pp; Japanese.	
The present sequence is a gene, which encodes a protein with vacuolar pH regulatory activities. The gene enables flower colour to be controlled via regulation of the vacuolar pH, colours can range from blue to red in colour spectrum. The gene is useful in controlling flower colour to give	

Example 7: page 45-49: 68pp: Japanese.

CC new breeds of colourful plants for cut flowers, particularly applicable
CC in horticulture.
XX
SQ Sequence 2553 BP; 666 A; 487 C; 535 G; 865 T; 0 other;

Query Match		37.7%;	Score 821.2;	DB 22;	Length 2553;
Best Local Similarity		70.3%;	Pred. No. 4.7e-174;		
Matches 1115;		Conservative 0;	Mismatches 468;	Indels 3;	Gaps 1;
QY	289	ttggattctctagtcgcaaacctgcttcgttatcgacatctgacacgctctgtggtt	348		
DB	543	tttggactctgtggaaagatgaacaacttaacaactctgacatcaatcagtggt	602		
QY	349	gggtgaactctcttggctgcaacttctgtgtattgttcttgcacatcttttgaa	408		
DB	603	tcggtaaacctgttctgtgcaactatttgcgctgtatttgcgctgtatttctgag	662		
QY	409	gagaatagatgatgaacgaatccatccgctctgttggattggctaggcactgggtt	468		
DB	663	gaaacagatgatgaatgagtcacactgacactgctgtgattggtgactggagtc	722		
QY	469	acatttttggattagtaaaagaaagctgcgactcttctgtctttagtgaagatctt	528		
DB	723	atcttctactaagtggaggaagaaactcacatattttagtgttcagcgaagatctt	782		
QY	529	ttcttcatactctttgcccacccattattcaatgcaggtttccaagttaaaaaagag	588		
DB	783	ttcttcatttacccttccacgacatcttttaagtctgggttccaggtgaaaaagaa	842		
QY	589	cagttttccgcaattctgactattatgtcttttggctgtgttggactattattct	648		
DB	843	tcatttccgcaattctgactattatgtcttttggctgtgttggactattattct	902		
QY	649	tgacaataatactctaggttaacacagttctttaagaagtggacattggaaccttt	708		
DB	903	ttcatattatatacgggtgtctatcttgacatttcaagaataatggatattggacactt	962		
QY	709	gacttgggtgattatctgtctattgttgcacatatttgcacagattcagtagtaca	768		
DB	963	gaaatggagatttaaccttgcattggagcaacttcttgcacagattctgtatgcacc	1022		
QY	769	ctgacaggttctgaatacaagacagacacacttgcctttacagctctgtattcggagaggt	828		
DB	1023	ttacaagtgcttaatacagaagaacacggttatgtacagctcagttgttggagaaggt	1082		
QY	829	gttgtgaatgatcaacgtcagttgttctcaacgctaticcagagctttgatctcact	888		
DB	1083	gttgtgaatgatccacatctgtagtctgttcaatgtctcagaaactttgacttatct	1142		
QY	889	cactaaaccacgaagctgcttttcatctcttcttggaactcttctgtattgtttctccta	948		
DB	1143	catatcagcagcgcaagctctgcataatgtgaactttctatactgtttgctgcgc	1202		
QY	949	agtaacttgggtgctgcaacggctctgataagtgctatgttatcaagaagctatcac	1008		
DB	1203	agcaacttctaggggtgtgttggcctactaagtgccctttataattagaactctcac	1262		
QY	1009	tttggaaaggcactcaactgaccgagaggttgcctcttatgatgcttattggcgttatcttct	1068		
DB	1263	tttggaaaggcactgactgactgaggttgcataatgatactcatgctgctacattatca	1322		
QY	1069	tatatgtctgagcttcttcgacttgagcgttatctcactgtgttttctctgtggtatt	1128		
DB	1323	tacatgtctggaattatctattttaaagtggaactcactggttttctctgtgggactc	1382		
QY	1129	gtcatgtccattacacatggcacaatgttaacggagagctcaagaatacaacaagaagcat	1188		
DB	1383	gtgatgtctcactatactctggcataatgtgactgagagctcaagagtcactaccaagcac	1442		
QY	1189	acctttgcaactttgtctatttctgcggagacatttatttctgtatgttggaaaggat	1248		
DB	1443	acgtttgctacattatctatttattgtcgaaataatctcatattctcttattgttggatggat	1502		

QY	1249	gccttgacattgacaaagtgagatccgtgagtgacacacccgggaacatcgatcgagtg	1308		
DB	1503	gccttgacattgagaagtggaagtttgLaagcgacagcccggaacatcaatgaagtc	1562		
QY	1309	agctcaacttaaatgggtctgtcatcgttggaaagacagcgttcttctccgttatcg	1368		
DB	1563	agctcaattctgtaggtctctgttttggttggaagggagccttgttttccctctgtca	1622		
QY	1369	tttctatcaactgaccgaagaatcaaacgagaaatacaacttttaacatgcaggtt	1428		
DB	1623	ttcttgcacacttgaccagaataaactcctgaggaacaagattagctttaaccagcaggtt	1682		
QY	1429	gtgatttgggtctgtctcatgagaggtgctgtatctatctatggctcttgcatacaacaag	1488		
DB	1683	acaatatgggtgggtggttgcgaggtgctgttctctatggccttgccttataatcag	1742		
QY	1489	tttacaaggccggcgacacagatgacgagggaatgcaatcatcatgacgagtagcgata	1548		
DB	1743	tttaccaggggaggtcatactcagttcacgtgcacaaTgcataaatgatacagagtagctatc	1802		
QY	1549	actgtctgtctttttagcacagtggttgggttatgctgaccaaaacc---actcataaagc	1605		
DB	1803	actgtgtccttttcagcacagtggttatttgggttgatgacaaaacactttaattctatta	1862		
QY	1606	tactattacgcacagaacgcacacagcagcagatgttattctgatacaacaccccaaaa	1665		
DB	1863	ttgtacccctcacaacacacttgatcagaatgatctctctgaaacgagtagctccaaa	1922		
QY	1666	tcatacatatccttcttggccaagactcgttcatgagccttcadgggaacccaact	1725		
DB	1923	tccttcatttgccacttcttgacagcacacaagactcagaagctgatctggtggcgacat	1982		
QY	1726	gtgctcgggctgacagtatacgttggcttcttgacacggcccactcgaaacgtgcattac	1785		
DB	1983	gtaccccgctcccaacagtttgcgagtgctctctgtcaaccccatctcacaacgttacattac	2042		
QY	1786	tactggagacaatttgatgactctctcatcgacccgcttcttgagagctgtggtttgta	1845		
DB	2043	tactgggagaaaatttgacaatgcatcattcagctcgtctgttttcggtgacagaggtttgta	2102		
QY	1846	ccctttgttccagtttctccaactga	1871		
DB	2103	ccctttgttccaggtacacactatga	2128		

RESULT 5
AAF75764
ID AAF75764 standard; DNA; 2423 BP.
XX
AC AAF75764;
XX
DT 14-MAY-2001 (first entry)
XX
DE Gene regulating the pH of vacuoles.
XX
KW Vacuole pH regulation; flower colour; ds.
XX
OS Petunia hybrida.
XX
PN WO200114560-A1.
XX
PD 01-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-JP05722.
XX
PR 24-AUG-1999; 99JP-02368000.
XX
PA (SUNR) SUNTORY LTD.
XX
PI Iida S, Tanaka S, Inagaki Y;
XX
DR WPI; 2001-191648/19.

Db	941	gtc	aagaa	ctactcc	aggcatgtattg	caattg	tctcttattg	ggagacaattca	1000
Qy	1226	t	t	t	t	t	t	t	1285
Db	1001	t	a	t	t	t	t	t	1060
Qy	1286	c	a	c	c	g	g	a	1345
Db	1061	g	e	t	t	g	g	g	1120
Qy	1346	c	a	c	g	t	c	t	1399
Db	1121	c	a	c	a	t	t	t	1180
Qy	1400	g	c	a	g	a	a	a	1459
Db	1181	a	c	g	t	c	a	t	1240
Qy	1460	c	t	g	t	a	t	a	1519
Db	1241	c	t	g	t	c	a	a	1300
Qy	1520	g	a	a	t	c	a	a	1579
Db	1301	t	g	a	t	g	c	c	1360
Qy	1580	g	t	a	t	g	t	a	1636
Db	1361	g	t	t	c	t	c	a	1420
Qy	1637	g	c	a	t	t	a	t	1696
Db	1421	g	a	a	t	g	a	a	1480
Qy	1697	c	g	t	t	a	t	g	1753
Db	1481	t	t	c	c	t	t	g	1540
Qy	1754	t	t	g	a	c	c	c	1813
Db	1541	t	g	a	t	g	a	a	1600
Qy	1814	t	g	c	a	c	c	t	1838
Db	1601	t	a	a	c	c	t	a	1625

RESULT 10

AAZ22611

ID AAZ22611 standard; DNA: 330 BP.

XX
XX

AAZ22611;

XX
DT 18-JAN-2000 (first entry)

XX Arabidopsis Na/H antiport transporter gene from EST clone T75860.

DE
XX

AA sodium: proton: antiport: transporter; salt tolerance; salt management

KW
transgeni

KW crop: SS.

XX
XX

OS Arabidopsis

XX

PN WO9947679-A2.

XX

PD 23-SEP-1999.

(APSE/) APSE M.
(SNED/) SNEDDEN W.
(AHAR) AHARON G.
Blumwald E, Apse M, Snedden W, Aharon G;
WPI; 1999-571840/48.
Nucleic acid molecules encoding sodium/proton transport polypeptides,
useful in genetic engineering salt tolerance in crop plants -

Sequence 330 BP. 92 A: 68 C: 66 G: 91 T: 13 other;

Query Match 11.68: Score 252.6; DB 20; Length 330;

Query Match	Best Local Similarity	Pred. No.	2.1e-47;
90.7%			

Best Local Similarity	30.1%	30.1%	30.1%
Matches	264	Conservative	27
Mismatches	0	Mismatches	27
Indels	0	Indels	0
Gaps	0	Gaps	0

1357 + + c c c + + a + c c + + c t a a c t t a c c c a a g a a a a a t c a a a c c g a g a a a a t c a a c t t t 1416

[illegible]

nb
1 ++++++tctatctaacttagccaaqaaqaatacaagcgagaaaaatcaaacttt 60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1417 aacatgcaagatttgtgatttggtggtctggtctcatgagaggctgtatctatggctctt 1476

Db 61 aacatgcagggtgtgtgatttgggtggtctggtctcatgagaggtgctgtatctctatggctctt 120

1526

QY 1477 gcatacaagaagttacaaggccgggcacacagatgtacgcgggaatgcaatcattgac

180

RESULT 11

AA7.22606

AAZ2206
ID AAZ22606 standard: DNA: 378 BP.

[illegible]

AAZ22606;

XX
XX

DT 18-JAN-2000 (first entry)

[illegible]

DE Rice Na/H antiport transporter gene from EST clone C91832.

XX

KW Sodium; proto;

KW transgenic pl

XX WO9947679-A2.
XX 23-SEP-1999.
XX 18-MAR-1999; 99WO-CA00219.
XX 18-MAR-1998; 98US-0078474.
XX 15-JAN-1999; 99US-0116111.
XX (BLUM/) BLUMWALD E.
XX (APSE/) APSE W.
XX (SNED/) SNEDDEN W.
XX (AHAR/) AHARON G.
XX Blumwald E, Apse M, Snedden W, Aharon G;
XX WPI; 1999-571840/48.
XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
XX useful in genetic engineering salt tolerance in crop plants -
XX Disclosure; Fig 8C; 93pp; English.
XX The invention relates to an isolated nucleic acid molecule encoding
XX a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
XX and capable of increasing salt tolerance in a cell. This sequence
XX corresponds to an Na/H antiport transporter gene fragment from the
XX rice expressed sequence tag (EST) library clone AAC91832. The Na/H
XX transporter polypeptides provide a means of intracellular salt
XX management, particularly in plants. The sequences are useful for
XX producing transgenic plants that are capable of surviving in soil
XX with high salt levels that would normally inhibit growth of the crop
XX species. This would be useful in farming land in areas that are
XX generally considered unproductive through salt accumulation and poor
XX irrigation, e.g. in India, Australia, and prairies in USA or Canada.
XX Commercial crops, such as potatoes, tomatoes, brassica, cotton,
XX sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
XX rye, barley, atriplex, sorghum, alfalfa, salicornia and others would
XX benefit from increased salt tolerance.
XX Sequence 378 BP; 88 A; 74 C; 89 G; 121 T; 6 other;
Query Match 10.4%; Score 236.4; DB 20; Length 378;
Best Local Similarity 77.2%; Pred. NO. 1.5e-41;
Matches 284; Conservative 0; Mismatches 82; Indels 2; Gaps 1;
QY 996 caagaagctatcttggagcactcaactgacgagaggttgcccttatgattcttat 1055
DB 1 caagaagctatcttggagcactcaactgacgagaggttgcccttatgattcttat 60
QY 1056 ggcgtatcttttatatctgtgagcttttcgacttgagcggtatcctcactgtgtt 1115
DB 61 ggcattaccttatatctgctgagctgtgtagatttgagcggtatcctcactgtatt 120
QY 1116 ttctgtgtattgtgattgctccattacatgacatgacatgtaacgagagctcaagaat 1175
DB 121 ctctctgtgtattgtaattgtcacattacacttggcatacgtccacagaggttcaagagt 180
QY 1176 aacaaacaaagcactcttggcactttctcattcttcgagagacattatttcttata 1235
DB 181 tacaacaaagcagcatttgcaactctctccttctcattgtgagactttctctctgta 240
QY 1236 ttttggaaatggatgcttgagcattgacaagtggagatccgtgagtcacacccgggaac 1295
DB 241 ttttgggaatggatgcttgagcattgaaaatggagatnccagtgacagacctgmaaa 300
QY 1296 atcgatcgacgtgagctcaactct--aatgggtctgtctatggttggaagagcagcttc 1353
DB 301 atccattngggtaagctcaacttttctgtagggattgtgttctctgattgggaagngctgcttt 360
QY 1354 gcttttcc 1361

DB 361 gnaattcc 368
RESULT 12
AAC95372
ID AAC95372 standard; cDNA; 1824 BP.
XX AAC95372;
XX 19-FEB-2001 (first entry)
XX Cat flea HMT Na/H transporter cDNA ORF, SEQ ID NO:1870.
XX Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;
XX flea infestation; vaccine; antiparasitic; therapeutic target;
XX diagnosis; detection; ss.
XX Ctenocephalides felis.
XX WO2000061621-A2.
XX 19-OCT-2000.
XX 07-APR-2000; 2000WO-US09437.
XX 09-APR-1999; 99US-0128704.
XX (HESK-) HESKA CORP.
XX Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;
XX WPI: 2000-656323/63.
XX P-PSDB; AAB29621.
XX Flea Malpighian tubule and head and nerve cord tissue derived nucleic
XX acids useful for the prevention, diagnosis and treatment of flea
XX infestations -
XX Claim 1; Page 894-895; 964pp; English.
XX The invention relates to novel cat flea (Ctenocephalides felis) nucleic
XX acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
XX or head and nerve cord (HNC) tissue. The invention also relates to the
XX encoded proteins. The invention additionally encompasses expression
XX constructs, recombinant viruses and recombinant cells comprising the
XX nucleic acids of the invention, recombinant production of the proteins,
XX antibodies against the proteins, a method of identifying inhibitors of
XX the proteins, and compositions comprising the nucleic acids, and the proteins they
XX administration to an animal. The nucleic acids, and the proteins they
XX encode may be used in the prevention, treatment and diagnosis of diseases
XX associated with flea infestations. For example, the nucleic acids may be
XX used to produce an HMT or HNC protein according to standard recombinant
XX DNA methodology by inserting the nucleic acids into a host cell and
XX culturing the cell to express the protein. The HMT and HNC nucleic acids
XX may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
XX and quantitate the presence of cat flea or other homologous nucleic acid
XX sequences in samples. They may also be used to study the expression and
XX function of the proteins and their role in metabolism. The HMT and HNC
XX proteins may be used as antigens in the production of specific
XX antibodies, and in assays to identify modulators (agonists and
XX antagonists) of HMT and/or HNC protein expression and activity. The
XX anti-HMT/HNC protein antibodies and antagonists may also be used to
XX downregulate protein expression and activity. The antibodies may also be
XX used as diagnostic agents for detecting the presence of flea polypeptides
XX in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
XX present sequence represents a cat flea HMT cDNA of the invention.
XX Sequence 1824 BP; 529 A; 330 C; 355 G; 610 T; 0 other;

Query Match 6.5%; Score 141.8; DB 21; Length 1824;
Best Local Similarity 49.7%; Pred. No. 1.8e-22;

Db 912 gcttcggtggtatcggtgtagtcttttgccttaattagtgctcttttggtaaacatggt 971
Qy 1009 ttgtgaagcactcaactgacgaaggttgcccttatgatgcttatgctatctttct 1068
Db 972 gatcttagaagtatccgtcccttagagttagtgatggttgggttacttatgcaacct 1031
Qy 1069 tatatgctgtagcttttgcacttgagcggtatccctcactggtttttctggtgatt 1128
Db 1032 tatgtttggcagaaggaattcattatcaggtataatggcgatattattcttggcatt 1091
Qy 1129 gtgatgtccattacacatgacgaatgtacgaagagctcaagaataacacaagaact 1188
Db 1092 gfgatgtccattacacacatttcaatttatcaacggttacacaataactatgcagcag 1151
Qy 1189 accttgcgaactttgatttcttcgcgagacattttttcttctgtatgttggaaatgat 1248
Db 1152 acgatgagaactttggtttttattgagaaactgtgtgttcttatttgggaatggct 1211
Qy 1249 gcttt 1253
Db 1212 atatt 1216

RESULT 15
AAC95371/C
ID AAC95371 standard; cDNA; 2080 BP.
XX
AC AAC95371;
XX
DT 19-FEB-2001 (first entry)
XX
DE Cat flea HMT Na/H transporter cDNA complement, SEQ ID NO:1869.
XX
KW Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;
KW flea infestation; vaccine; antiparasitic; therapeutic target;
KW diagnosis; detection; ss.
XX
OS Ctenocephalides felis.
XX
PN WO200061621-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US09437.
XX
PR 09-APR-1999; 99US-0128704.
XX
PA (HESK-) HESKA CORP.
XX
PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;
XX
XX WPI; 2000-656323/63.
DR P-PSDB; AAB29621.
XX
XX Flea Malpighian tubule and head and nerve cord tissue derived nucleic
PT acids useful for the prevention, diagnosis and treatment of flea
PT infestations -
XX
PS Claim 1; Page 893-894; 964pp; English.
XX
XX The invention relates to novel cat flea (Ctenocephalides felis) nucleic
CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
CC or head and nerve cord (HNC) tissue. The invention also relates to the
CC encoded proteins. The invention additionally encompasses expression
CC constructs, recombinant viruses and recombinant cells comprising the
CC nucleic acids of the invention, recombinant production of the proteins,
CC antibodies against the proteins, a method of identifying inhibitors of
CC the proteins, and compositions comprising the inhibitors for
CC administration to an animal. The nucleic acids, and the proteins they
CC encode may be used in the prevention, treatment and diagnosis of diseases
CC associated with flea infestations. For example, the nucleic acids may be
CC used to produce an HMT or HNC protein according to standard recombinant
CC DNA methodology by inserting the nucleic acids into a host cell and

CC culturing the cell to express the protein. The HMT and HNC nucleic acids
CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC and quantitate the presence of cat flea or other homologous nucleic acid
CC sequences in samples. They may also be used to study the expression and
CC function of the proteins and their role in metabolism. The HMT and HNC
CC proteins may be used as antigens in the production of specific
CC antibodies, and in assays to identify modulators (agonists and
CC antagonists) of HMT and/or HNC protein expression and activity. The
CC anti-HMT/HNC protein antibodies and antagonists may also be used to
CC downregulate protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of flea polypeptides
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC present sequence represents a cat flea HMT cDNA of the invention.
XX
SQ Sequence 2080 BP; 703 A; 389 C; 349 G; 639 T; 0 other;

Query Match 6.5%; Score 141.8; DB 21; Length 2080;
Best Local Similarity 49.7%; Pred. No. 1.8e-22;
Matches 390; Conservative 0; Mismatches 392; Indels 3; Gaps 1;

Qy 472 attttgttattagtaaaaggaagaaagctgcctcttctctgtctttagtaagatcttttc 531
Db 1649 ATGTCGCTAAAATATTTGCAAAATTGGAAGAAATGAGAAGCCCTTTTACCACAGCGTTT 1590
Qy 532 ttcatatatcttttgcacccatttatattcaatgcagggtttccaagtaaaaaagaagcag 591
Db 1589 TTCTTAGTGTCTTACGCGCTATAATATTTGAATCCGGGTATAATTTGCATAAAGGTAAT 1530
Qy 592 tttttcgcgaattcgtgactattatgcttttttgggtgctgttggactattattcttgc 651
Db 1529 TTTTTCAAAATATTTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1470
Qy 652 acaatcatatctctaggtgaacacagcttctttaaagaagttggacatttggaaacctttgac 711
Db 1469 TTTCCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1410
Qy 712 ttgggtgattatcttctgttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc 771
Db 1409 TTTTGTGAATCCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCT 1350
Qy 772 caggtctctgaa---tcaagacagacacaccttcttcaagacatttgcatttgcatttgcatttgc 828
Db 1349 GCTATTTTCCATGCTTTAGACGTGGACCCAGTTTAAACATGTTGGTGTTCGGAGAAAGT 1290
Qy 829 gttgtgaatgatgaacgctcagttgt 888
Db 1289 ATTTTAAATGATGCTATTTCAATTTGTTTAAACAACCTGCAGTTTGGAAATCCAACAATCCT 1230
Qy 889 cacctaaacacgaagctgttttctcattctcttggaaaccttctgtatttctgtttcttctccta 948
Db 1229 TTAATGACGACTGCTGAAGCTGTAGTCTCCGGTTTAAATAGGTTTGTAAATGTTTCTTTT 1170
Qy 949 agtaccttgccttgcctgaacccggtctgataagtcgctgataagtcgctgataagtcgctgata 1008
Db 1169 GCTTCGGCTGATCGGTGTAGCTTTGCCCTTAATYAGTGTCTCTTTTGTGGAACATGTT 1110
Qy 1009 ttgtgaaggcaactcaactgacccgagaggttgccttatgatgtcttatggttgccttatggttgc 1068
Db 1109 GATCTTAGAAAGTATCCGCTCCTTAGAGTATAGTATGTTGGTGTGTTTACTTATGACCT 1050
Qy 1069 tatatgcttgcctgagcttttcgcacttgagcgggtatctcactgtgttttctgtgtatt 1128
Db 1049 TATCTTTGGCAGAGAAATTCATTTATCAGGTATAATAGGCGATATTATTCTGTGGCAAT 990
Qy 1129 gtgatgtccattacacatggaacaaatgtacggagagctcaagaataacacaagaagcat 1188
Db 989 GTGATGTCCTTACACATTTCAATTTATCAACGGTTACACAAATACTATGCAGCAG 930
Qy 1189 accttgcgaactttgcttatttcttgcggagacatttatttcttctgtatgttggaaatggat 1248
Db 929 ACGATGAGAACTTTGGCTTTTATTTCGAGAAACTTGTGCTGTTGCTTATTATTAGGAATGGCT 870

Qy 1249 gcctt 1253
||
Db 869 ATATT 865

Search completed: November 18, 2001, 11:08:22
Job time: 4566 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 09:49:01 ; Search time 81 seconds

(without alignments)
5089.742 Million cell updates/sec

Title: US-09-271-584A-1

Perfect score: 2178
Sequence: 1 cctctctgttcttcctcg.....aaaaaaaaaaaaaaaaaaaaa 2178

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53.8	2.5	2007	3	US-08-747-221B-36
2	53.8	2.5	2007	3	US-08-747-221B-38
3	53.8	2.5	2007	4	US-09-005-051-36
4	53.8	2.5	2007	4	US-09-005-051-38
5	48.8	2.2	1169	3	US-09-100-391-5
6	48.8	2.2	1302	3	US-09-100-391-11
7	48.8	2.2	2643	3	US-09-100-391-9
8	46.6	2.1	5852	1	US-07-867-106-2
9	46	2.1	3138	1	US-07-867-106-4
10	45.4	2.1	1472	4	US-08-781-420-10
11	45.4	2.1	1472	4	US-08-781-420-12
12	44.4	2.0	1368	3	US-08-874-563-5
13	44.4	2.0	1368	3	US-08-577-483-14
14	44.4	2.0	4253	3	US-08-577-483-7
15	44	2.0	6152	4	US-08-973-462-1
16	43.8	2.0	3871	2	US-08-599-455B-3
17	43.8	2.0	3871	4	US-09-069-781B-3
18	43.6	2.0	961	6	5194596-16
19	43.6	2.0	961	6	5219739-16
20	43.4	2.0	1273	1	US-08-507-431-1
21	43.4	2.0	1273	3	US-08-902-655A-1
22	43.4	2.0	1273	3	US-09-116-622-1
23	43.4	2.0	1273	3	US-09-219-277-1
24	43.4	2.0	1273	4	US-09-599-661-1
25	43.4	2.0	1393	1	US-08-174-467-18
26	43.4	2.0	1393	3	US-08-452-071-18
27	42.8	2.0	4254	2	US-08-443-639-7

28 42.6 2.0 654 5 PCT-US95-06406A-11
29 42.6 2.0 790 6 5194596-8
30 42.6 2.0 1493 6 5340934-5
31 42.6 2.0 3157 6 5198347-3
32 42.6 2.0 7218 1 US-08-232-463-14
33 42 1.9 5852 1 US-07-867-106-2
34 41.4 1.9 114 1 US-08-120-827-99
35 41.4 1.9 114 1 US-08-478-675-99
36 41.4 1.9 1279 4 US-09-277-716-31
37 41.2 1.9 2230 4 US-08-378-313-24
38 41 1.9 1939 1 US-07-715-751B-2
39 41 1.9 2133 4 US-09-187-124-1
40 40.8 1.9 1264 2 US-08-758-621-13
41 40.8 1.9 1264 4 US-09-107-858-13
42 40.8 1.9 1357 6 5340934-7
43 40.6 1.9 2444 3 US-08-906-791-1
44 40.6 1.9 11517 1 US-07-920-281C-1
45 40.6 1.9 11517 4 US-08-466-277-1

ALIGNMENTS

RESULT 1
US-08-747-221B-36
; Sequence 36, Application US/08747221B
; Patent No. 6063610
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary W.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: No. 6063610el Carboxylesterase Nucleic Acid
; TITLE OF INVENTION: Molecules, Proteins and Uses Thereof
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,221B
; FILING DATE: No. 6063610ember 12, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: FC-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2007 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..1594
; US-08-747-221B-36

Query Match 2.5%; Score 53.8; DB 3; Length 2007;
Best Local Similarity 52.4%; Pred. No. 0.00029;
Matches 116; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy 2134 ttTggttttgcTcgaaaaaa 2178

; ORGANISM: HOMO SAPIENS
 US-09-100-391-11

```

Query Match          2.2%; Score 48.8; DB 3; Length 2643;
Best Local Similarity 70.7%; Pred. No. 0.0057;
Matches 65; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 2087 tattgtttgtgtaacacaaactacacattgtttatgttttgaattgggttttgt 2146
      ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2546 tttttattttactgttataataattattaaacttctctgttaataagaataaagtgttact 2605
      ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2147 tcgaaaaaataaaaaaataaaaaaataaaaaa 2178
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2606 tggaaaaaataaaaaaataaaaaaataaaaaa 2637
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT      8
US-07-867-106-2
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA

```


Db 195 TTTTATATTTTATAAATTTATATAAATCAATATATAAATCAGAAAAATACATAAAATG 254
QY 2024 tgagagcaaaacaaacacttggcaacttgaagtggttgattgatgtatgtaattatat 2083
Db 255 TGAAGATTAAAGAAATATAGAAATTAATATATATATATATATATATATATATATAT 314
QY 2084 tcattattgtttgtgaacaaactacacatttggttatgttt 2129
Db 315 ATATATATATATATATATATATATAAATGTATGTGTGTGTGTGT 360

RESULT 14
US-08-577-483-7
; Sequence 7, Application US/08577483
; Patent No. 6100451
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: Yin, Shaohui
; APPLICANT: Cornett, Catherine A.G.
; TITLE OF INVENTION: Transcriptional Control Sequences and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/577,483
; FILING DATE: 22-DEC-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 32,164
; REFERENCE/DOCKET NUMBER: 07678/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4253 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1216..1327
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1454..1718
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1805..2182
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2259..2477
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2609..2747
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 2902..3148
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3261..3555
US-08-577-483-7

Query Match 2.0%; Score 44.4; DB 3; Length 4253;
Best Local Similarity 54.2%; Pred. No. 0.084;
Matches 90; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1964 ttgtgtaaatatccatttgaataattgttggagagacagaaatctgtcctaacttt 2023
Db 195 TTTTATATTTTATAAATTTATAAATCAATATATAAATCAGAAAAATACATAAAATG 254
QY 2024 tgagagcagaagcaaacatggcaacttgaagtggttggattgatgtatgtaattatat 2083
Db 255 TGAAGATTAAAGAAATTTATAGAAATTAATATATATATATATATATATATATATATAT 314
QY 2084 tcattattgtttgtgaacaaactacacatttggttatgttt 2129
Db 315 ATATATATATATATATATATATATAAATGTATGTGTGTGTGTGTGT 360

RESULT 15
US-08-973-462-1/c
; Sequence 1, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DAUBERSIES, PIERRE
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 6152
; TYPE: DNA
; ORGANISM: P. falciparum
US-08-973-462-1

Query Match 2.0%; Score 44; DB 4; Length 6152;
Best Local Similarity 52.8%; Pred. No. 0.12;
Matches 95; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1923 ttgtatttttttgtagaaaagggtgattcaaatgatgctttgtgtaattatccatt 1982
Db 6128 TTTTATATAGGAAAAATATTTATATAAATTTAAAAATTTAAAAATTTAAAAATTTAAACATC 6069
QY 1983 tgtaaatattgttgtagagacagaaatctgtcctaacttttgagagacagaaagcaaac 2042
Db 6068 GTTTATATGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 6009
QY 2043 atggcaacttgaagtggttgattgatgtatgtaattatcatatttggttgtga 2102
Db 6008 AAGCTAAATTTTATCTTCTAATA 5949

Search completed: November 18, 2001, 11:05:34
Job time: 4593 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 18, 2001, 08:39:10 ; Search time 1652.68 Seconds
(without alignments)
12457.533 Million cell updates/sec

Title: US-09-271-584A-1
Perfect score: 2178
Sequence: 1 cctctgtttctgttctctg.....aaaaaaaaaaaaaaaaaaaa 2178

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	EST:*
-	1: gb_est1:*
	2: gb_est2:*
	3: gb_est3:*
	4: gb_est4:*
	5: gb_est5:*
	6: gb_est6:*
	7: gb_est7:*
	8: gb_est8:*
	9: gb_est9:*
	10: gb_est10:*
	11: gb_est11:*
	12: gb_est12:*
	13: gb_est13:*
	14: gb_est14:*
	15: gb_est15:*
	16: gb_est16:*
	17: gb_est17:*
	18: gb_est18:*
	19: gb_est19:*
	20: gb_est20:*
	21: gb_est21:*
	22: gb_est22:*
	23: gb_est23:*
	24: gb_est24:*
	25: gb_est33:*
	26: gb_est34:*
	27: gb_est35:*
	28: gb_est36:*
	29: gb_est37:*
	30: gb_est38:*
	31: gb_est39:*
	32: gb_est40:*
	33: em_estba:*
	34: em_estfun:*
	35: em_esthum1:*
	36: em_esthum2:*
	37: em_esthum3:*
	38: em_esthum4:*
	39: em_esthum5:*
	40: em_esthum6:*
	41: em_esthum7:*
	42: em_esthum8:*
	43: em_esthum9:*
	44: em_esthum10:*
	45: em_esthum11:*
	46: em_esthum12:*
	47: em_esthum13:*
	48: em_esthum14:*
	49: em_esthum15:*
	50: em_esthum16:*
	51: em_esthum17:*
	52: em_esthum18:*
	53: em_esthum19:*
	54: em_esthum20:*
	55: em_esthum21:*
	56: em_esthum22:*
	57: em_esthum23:*
	58: em_esthum24:*
	59: em_esthum25:*
	60: em_esthum26:*
	61: em_esthum27:*
	62: em_esthum28:*
	63: em_estin1:*
	64: em_estin2:*
	65: em_estin3:*
	66: em_estin4:*
	67: em_estin5:*
	68: em_estom1:*
	69: em_estom2:*
	70: em_estov1:*
	71: em_estov2:*
	72: em_estpl1:*
	73: em_estpl2:*
	74: em_estpl3:*
	75: em_estpl4:*
	76: em_estpl5:*
	77: em_estpl6:*
	78: em_estpl7:*
	79: em_estpl8:*
	80: em_estpl9:*
	81: em_estpl10:*
	82: em_estro1:*
	83: em_estro2:*
	84: em_estro3:*
	85: em_estro4:*
	86: em_estro5:*
	87: em_estro6:*
	88: em_estro7:*
	89: em_estro8:*
	90: em_estro9:*
	91: em_estro10:*
	92: em_estro11:*
	93: em_estro12:*
	94: em_estro13:*
	95: em_estro14:*
	96: em_estro15:*
	97: em_estro16:*
	98: em_estro17:*
	99: em_estro18:*
	100: em_estro19:*
	101: em_estro20:*
	102: gb_est25:*
	103: gb_est26:*
	104: gb_est27:*
	105: gb_est28:*
	106: gb_est29:*
	107: gb_est30:*
	108: gb_est31:*
	109: gb_est32:*
	110: gb_est41:*
	111: gb_est42:*
	112: gb_est43:*
	113: gb_est44:*
	114: gb_est45:*
	115: gb_est46:*
	116: gb_est47:*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est167:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*
```

```
190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

BASE COUNT	119 a	95 c	112 g	176 t	randy phenomaker:
ORIGIN					
Query Match	13.7%	Score 297.8;	DB 167;	Length 502;	
Best Local Similarity	74.7%;	Prod. No. 1.8e-60;			
Matches 374;	Conservative	0;	Mismatches 127;	Indels 0;	Gaps 0;
620	tttttgctgttgggaactattattctttgcaaatcatctctgaagtgttaacacagt	679			
Ddb	2 TGTGTGGTCAATTATGGACATTAATATCTTGTACCAATATAACTTTGGTGGCCACACAAA	61			
680	tettaagaagtgtggacatttgaacacctttgacttgggtgatttatcttctatttggtgcca	739			
Ddb	62 TTTTAAAGAGGTTGGATGTGTGGTCCACTGGAAATTAGGGATTTCCTAGCAATTGGTGCAA	121			
740	tatttctgcaacagattcagtagtatgacatcgcagttctgaatcaaacacagacacct	799			
Ddb	122 TATTGTGTCGAACGATTCGTGTGGACATTCGAGTGTCAATCAGATGAGACCCCTT	181			
800	tgcttaacagtcctgtatctcgagagggtgtgtgaatgatgcaacgtcagtttgggtct	859			
Ddb	182 TGCTGTACAGTCTGTGTAATTTGGGAGGGTGTGTGAATGATGCTACATCAGTGGTGCTTT	241			
860	tcaacgggattcagagcttctgactcactcacctaacccagcaagctctttcatcttc	919			
Ddb	242 TCATATGCATCCAAAGCTTTGACCTCAACCAATTCAGCCCTTCAATTGCTGGGCACTTTT	301			
920	ttggaacactctgtatttcttctcttaagtaccttggttggtgctgcacacccggtctga	979			
Ddb	302 TGGGAATATTCCTGTATCTAATTATTGCAAGCAACAATGCTTGGAGTGTGTGACAGGTCTAC	361			
980	taagtgcgtatgttatacaagaagctatacttftggaaggcaactcaactgaccgagaggtt	1039			
Ddb	362 TTAGTGCCTTACATTTATTAAGAAGCTGTACATTTGGCAGGCACCTTACAGATCGCAGGTTG	421			
1040	cccttatgatgcttatggctatctttctttatatcttgctagagcttttcgacttgagcg	1099			
Ddb	422 CTCCTATGATGTTAATGGCAATCTCTCTACATGCTTCTGAATGCTCTATCTTCTGAGTG	481			
1100	gtactctcaactgtgtttttct	1120			
Ddb	482 GCATCTCTACGTATCTTTT	502			

RESULT	4				
A1938253					
LOCUS	A1938253	486 bp	mRNA	EST	02-AUG-1999
DEFINITION	sc41102.y1 Gm-cl014 Glycine max cDNA clone				
	Gm-cl014-1708 5' similar to TR:004655 004655 SIMILAR TO				
	SODIUM/HYDROGEN EXCHANGER. ; mRNA sequence.				

ACCESSION	A1938253	
RELEASED	A1938253	
PROTEIN	A1938253.1	GI:5677123
KEYWORDS	EST	
SOURCE	soybean.	
ORGANISM	Glycine max	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.	

REFERENCE
AUTHORS

1 (bases 1 to 486)

Shoemaker R., Keim, P., Vodkin, L., Expelding, J., Corvelli, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST project

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Teacher Education on Classroom Practice	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Importance of Teacher Collaboration	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Effect of Teacher Attitudes on Student Behavior	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the School in the Development of the Child	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of Teacher Training on Student Learning	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63110
Tel: 314 286 1800
Fax: 314 286 1810

Fax: 314 200 1010
Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Possible reversed clone: similarity on wrong strand

Seq primer: -40RP from Gibco
High quality sequence stop: 381.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS"
/clone_lib="Gm-cl014"
/tissue_type="Leaves, 2
grown"
/lab host="nml0108"

```

/lab_host="DH10B"
 Vector: p77T3pac (p77T3, Pharmacia); Site_1: EcoRI;
 Site_2: HindIII; This cDNA library was constructed from
 mRNA isolated from leaves of 2-3 week old greenhouse grown
 plants. Complementary DNA was synthesized from mRNA using
 a 3' anchored poly (dT) primer. EcoRI adapters were
 ligated to the blunt-ended cDNA fragments followed by
 digestion with EcoRI and HindIII. The cDNA fragments were
 directionally cloned into the EcoRI-HindIII restriction
 site of the p77T3-pac vector. The ligated cDNA fragments
 were transformed into DH10B host cells (Gibco BRL). This
 library was constructed by Dr. Randy Shoemaker and Dr.
 John Erpelzing.

BASE COUNT	122 a	101 c	99 g	164 t
ORIGIN	JOMI EXPEDING.			

```
Query Match      13.4%; Score 292.8; DB 103; Length 486;
Best Local Similarity 75.6%; Pred. No. 2.7e-59;
Matches 363: Conservative 0; Mismatches 117; Indels 0; Gaps 0;
```

[illegible]

1225	atttctctgtatgtgtggaatggatgcctctggacaattgacaagttggagatccgtgaatgac	1284
Oy		
127	ATCTTCCCTTTATGTGTGGTATGGATGCGCTTGGACATTTGAAAAATTTGTCATGAT	186
Db		
1285	acaccgggaacatcgatcgcgagtgagctcaatcctaaggctgtgtgcattggttgaaaga	1344
Oy		
187	AGCCCTTGGAAACATCTGTAGCAACTAGTTTCAGTATTTTGGGTCTAATCTCTCTTGAAGA	246
Db		

[illegible]

QY	943	ctcctaagtaccttgccttggctgctgcacacgggtctgataagtgccgtatgttatcaagaag	1002
Db	301	ATCGGAAGCACATTGCTTGGGGTTTTTCACAGGTCTTCTCAGTGCATTACATTAATAAG	360
QY	1003	ctatacttctgaaagcactcaactgacgcagaggttgccttatgatcttatgcttat	1062
Db	361	CTATACATTGGCAGGCACCTACAGATCGTGAGGTGGCTCTTATGATGCTTAATGGCATAC	420
QY	1063	cttt 1066	
Db	421	cttt 424	
RESULT	6		
LOCUS	AW685820		
DEFINITION	AW685820	553 bp mRNA EST 15-JUN-2000	
ACCESSION	NF035E04NR	5'	Medicago truncatula cDNA clone
VERSION	AW685820		
KEYWORDS	AW685820.1	GI:7560556	
SOURCE	EST.		
ORGANISM	barrel medic.		
REFERENCE	Medicago truncatula		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Medicago. 1 (bases 1 to 653)		
TITLE	Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May, G.D. and Paiva,N.L.		
JOURNAL	Expressed Sequence Tags from the Samuel Roberts Noble Foundation		
COMMENT	Medicago truncatula nodulated root library		
FEATURES	Source		
BASE COUNT	167 a 133 c 148 g 205 t		
ORIGIN			
Query Match	11.68;	Score 253, DB 119;	Length 653;
Best Local Similarity	62.8;	Prod. No. 8.6e-50;	
Matches 410;	Conservative 0;	Mismatches 240;	Indels 3; Gaps 1;
QY	1203	gtcattcttgcggagacatttattcttctgtattgttgaaaggatgccttgacattga	1262
Db	1	GTCAATTCATTGCTGAGATATTTATCTTCCTTGTGTGGATGGATGCACTAGATATGGA	60
QY	1263	caagttggagatcgtgaatgacacacgggaacatcgatcgacgtgaactcaatccta	1322
Db	61	GAAGTGGCGATTTCGTAAGTCAAAAGCCAAAAAATCAATTTGGGTGATGCTTCTT	120
QY	1323	gggtctggttcattggttgaaagacagcaggttcgtcttcttcgttatcggtttctatc	1382

/note=Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRU2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA. "

BASE COUNT	92 a	58 c	66 g	91 t	13 others
ORIGIN					
Query Match	11.6%	Score 252.6;	DB 189;	Length 330;	
Best Local Similarity	90.7%;	Pred. No. 9.7e-50;			
Matches 264;	Conservative	0;	Mismatches 27;	Indels 0;	Gaps 0;
QY	1357	tttcggtatcggtttctatctaacttagccaagaataccaagcgagaaatacaacttt	1416		
Db	1	TTTTCCGTTATCGTTTCATCTAACTTAGCCACAAGAGATCAAGAAGCGAGAAATCAACCTT	60		
QY	1417	acatgcaggttgtagattggtggtctgctcatgagagggtcgtgatatctatgagcttt	1476		
Db	61	AACATGCAAGTGTGTGATTTGGTGTCTGGTCTCATGAGAGGTGCTGTATCTATGAGTCTTT	120		
QY	1477	gcatacaaaagtttacaagggcggggcacacagatgtacgcgggaatgcaatcatgatc	1536		
Db	121	GCATACAAAGTTTACAAGGGCGGGCACACAGATGTACGNGGGAATGCAATCATGATC	180		
QY	1537	acgagtagcataactctctcttttttagcacagtggtttgggtatctgtagcaaaacca	1596		
Db	181	ACGNTAGCAGTAACTGCTCTGTTTTTATGACACASTGGTGTGGTATGCTGACCAACCA	240		
QY	1597	ctcataactcactctatccgcgcagcaagcgcacacacagatgtatct	1647		
Db	241	NTCATAACTACTATTATTACCGNACCANGAACCGTCATCAACGNGGCATGT	291		
RESULT	8				
LOCUS	BE643915	469 bp	mRNA	EST	31-AUG-2000
DEFINITION	NXCI_048_H12_F NXCI (Nsf xylem Compression wood Inclined) Pinus taeda cDNA clone NXCI_048_H12_5', mRNA sequence.				
ACCESSION	BE643915				
VERSION	BE643915.1	GI:9956522			
KEYWORDS	EST.				
SOURCE	loblolly pine.				
ORGANISM	Pinus taeda				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 469)				
AUTHORS	Sederoff, R.				
TITLE	Molecular Basis of Wood Formation in the Pine Megagenome				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Johnson, Arthur North Carolina State University Tel: 919 515 7800 Fax: 919 515 7801 Email: ajohnson@unity.ncsu.edu Seq primer: T3.				
FEATURES	Location/Qualifiers				
source	1..469				
	/organism="Pinus taeda"				
	/strain="Coastal plain loblolly pine from North Carolina"				
	/db_xref="taxon:3352"				
	/clone_lib="NXCI_048_H12"				
	/tissue_type="Xylem"				
	/cell_type="Compression"				
	/dev_stage="Juvenile"				
	/lab_host="XLI-Blue"				
	/note="Vector: Bluescript SK; Site 1: Eco RI; Site 2: XhoI"				

```

FEATURES
  source
    location/Qualifiers
      1..378
        /organism="Oryza sativa"
        /strain="Nipponbare"
        /db_xref="taxon:4530"
        /clone="E31686_1A"
        /clone_lib="rice panicle"
        /dev_stage="shorter than
        /host="Oryza sativa"

```

BASE COUNT	88 a	74 c	89 q	121 t	6 others
noce- organ. panicle					

Query Match 10.4%; Score 226.4; DB 156; Length 378;
Best Local Similarity 77.2%; Pred. NO. 1.7e-43;

996 caaagaactatactttggaaagccactcaactgaccgagaggttgcccttatgatgcttat 1055

[illegible][illegible]

UD8 181 TACAACAAAGCAGGCATTTCGAATCTGACATTCATTCGTCGAGACCTTTCCTCTCCGCTA 240

Db 241 TGTGGGATGGATGTCATTGGATATTGAAAAATGGGAGNTTNCAGTGACAGACCTGGNAA 300

Db 301 ATCCATTNGGGTAAGCTCAATTGTGCTAGGGATTGGTTCCCTGCTGCTTT 360

Db 361 GNAATTCC 368

REF ID: A1993308

CDNA CLONE 701430013, mRNA sequence.
ACCESSION AT993308

SOURCE
thale cress.
arabidopsis thaliana

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan

TITLE Arabidopsis thaliana Gene Expression MicroArray

ACCESSION A1941126
 VERSION A1941126.1 GI:5688111
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
 1 (bases 1 to 352)

```

QY 923 gaaactcttctgtattgtttctcttaagtaacctctgttggtg 964
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 GAAATTTCTGTATCTATTATTATGCAAGCAACATGCTTGAG 344

RESULT 14
AW691011 654 bp mRNA EST 15-JUN-2000
LOCUS NF040D08ST1F1000 Developing stem Medicago truncatula cDNA clone
DEFINITION NF040D08ST 5', mRNA sequence.
ACCESSION AW691011
VERSION AW691011.1 GI:7565747
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 654)
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
,R.A.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula stem library
JOURNAL Unpublished (2000)
COMMENT Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 654 Std Error: 0.00
Plate: 040 row: D column: 08
Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES
source
1..654
Location/Qualifiers
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF040D08ST"
/clone_lib="Developing stem"
/tissue_type="stem"
/dev_stages="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of
internodal stem segments"
BASE COUNT 186 a 122 c 122 g 224 t
ORIGIN

Query Match 8.9%; Score 194.6; DB 119; Length 654;
Best Local Similarity 76.9%; Pred. No. 7.1e-36;
Matches 250; Conservative 0; Mismatches 74; Indels 1; Gaps 1;

QY 294 tctctagtcgaaactgccttggtatcgatcgtatcgatcgtatcggttgcgtt 353
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 323 TCTATTGTTCAAAACATATCAATGTTATCCACTTCGATCATGCTCTGTGTTCTAT 382

QY 354 gaactcttctgtgcactctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 413
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 383 GAACCTGTTGTGGGACACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 442

QY 414 tagatggtatgaacgaatccatccacgcctgtgtgtgtgtgtgtgtgtgtgtgt 473
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 443 TCGCTGGATGAATGAATCCATCACTGCCTCTTTGATTGGTATATGCATGCTGTGTATGAT 502

QY 474 ttgtgttagtagtaaaagaaacgcgcgtcattctctctgtctctgtctctgtctgt 533
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 503 TTTCGTGTTTATGTTGGTGAAGAAAGTCCCATATTTCTGTTTTCAGTGAAGATCTTTTCTT 562

QY 534 catatctcttttgcaccatttatattcaatg-cagggtttcaagtataaaagaagcagt 592
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 563 TATATACCTGTCGCCCAATATATATTCATGTCGCCGGGTTTCAAGTAAAGAAAGCCATG 622

```

```

QY 593 ttctccgcaatttcgtgactattat 617
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 623 TTTTTCGTCAACTGTTTCATGACTAT 647

RESULT 15
AW694621 393 bp mRNA EST 20-DEC-2000
LOCUS NF078C11ST1F1085 Developing stem Medicago truncatula cDNA clone
DEFINITION NF078C11ST 5', mRNA sequence.
ACCESSION AW694621
VERSION AW694621.2 GI:11934170
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 393)
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
,R.A.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula stem library
JOURNAL Unpublished (2000)
COMMENT On Apr 14, 2000 this sequence version replaced gi:7569383.
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 757 Std Error: 0.00
Plate: 078 row: C column: 11
Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES
source
1..393
Location/Qualifiers
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF078C11ST"
/clone_lib="Developing stem"
/tissue_type="stem"
/dev_stages="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of
internodal stem segments"
BASE COUNT 111 a 94 c 72 g 116 t
ORIGIN

Query Match 8.8%; Score 192.6; DB 119; Length 393;
Best Local Similarity 71.0%; Pred. No. 2e-35;
Matches 255; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 1288 ccgggaacatcgatcgagtcagtcacttaacttaacttaacttaacttaacttaacttaact 1347
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 CCTGGAACATCTATAGCAGCAAGTTCAGTATTCTTGGGCTATACTCTTCTGGAAGACA 64

QY 1348 gcttcgtcttcctcggtattgttggttcgttcgttcgttcgttcgttcgttcgttcgt 1407
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 GCGTTTGTTCCTTCCTTATCTCTTATCCACTTGACTTAAAAAATCACAAATCAGAAG 124

QY 1408 atcaactttaacatcgagtttgatttggttcgttcgttcgttcgttcgttcgttcgttcgt 1467
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 125 ATCTCCTTCAGCAGCAAGTGATCATTTGGTGGGCTGGTCTTATGAGAGGTGCTGTTTCA 184

QY 1468 atggtctcttcatacaacaagtttaacaagcgcgagcacacagatgtacgcgggaatgca 1527
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 185 ATGGCACTTGGCTATATATCAGTTCCACATGTCGGGGCATACTCACTGCGTAGCAATGCA 244

QY 1528 atcatgatacagtagtataactgtctgtctcttttagcacagtggtgttggtatgtg 1587
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 245 ATCATGATAACCAAGCACCACACTGTTGCTCTTTTCAGCACAGTGTGTTGGTTCGTCG 304

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:26:37 ; Search time 15.28 Seconds
(without alignments)
792.328 Million cell updates/sec

Title: US-09-271-584A-2

Perfect score: 2755
Sequence: 1 MLDSLVKLSLSTSDHASV.....FVFPVGSPTERNPPDLSKA 538

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*

1: /cgn2.6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2.6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2.6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2.6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2.6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	471	17.1	832	2	US-08-677-734A-12
2	467	17.0	831	2	US-08-677-734A-11
3	462	16.8	834	2	US-08-677-734A-9
4	462	16.8	834	2	US-08-677-734A-10
5	104	3.8	429	2	US-08-677-049-5
6	99	3.6	462	3	US-09-238-796-2
7	98.5	3.6	397	1	US-08-098-141-2
8	95	3.4	508	1	US-08-472-028A-4
9	95	3.4	508	2	US-08-808-931-4
10	95	3.4	508	3	US-08-808-323-4
11	95	3.4	508	3	US-09-050-603A-4
12	95	3.4	508	3	US-09-102-420B-4
13	95	3.4	508	4	US-09-071-296-4
14	95	3.4	508	4	US-09-196-268-4
15	95	3.4	508	4	US-09-015-683-4
16	95	3.4	508	4	US-09-191-998-4
17	93	3.4	342	3	US-08-852-824-2
18	93	3.4	694	2	US-08-895-522-4
19	93	3.4	694	3	US-09-193-391-4
20	92.5	3.4	649	4	US-08-800-291B-5
21	92.5	3.4	650	4	US-08-800-291B-4
22	92.5	3.4	1711	3	US-08-369-822C-10
23	92.5	3.4	1711	3	US-08-582-776C-10
24	92.5	3.4	1711	3	US-08-434-831B-10
25	91.5	3.3	325	1	US-08-118-270-51
26	91.5	3.3	325	5	PCR-US93-08528-51
27	91.5	3.3	649	4	US-08-800-291B-6

28 90.5 3.3 635 2 US-09-014-969-11 Sequence 11, Appl
29 90 3.3 359 1 US-08-148-209A-4 Sequence 4, Appl
30 89.5 3.2 1059 4 US-09-134-513-2 Sequence 2, Appl
31 88.5 3.2 341 2 US-08-846-762-92 Sequence 92, Appl
32 88.5 3.2 471 1 US-07-996-772A-11 Sequence 11, Appl
33 88.5 3.2 471 4 US-09-032-742-2 Sequence 2, Appl
34 88.5 3.2 616 1 US-08-149-100-2 Sequence 2, Appl
35 88 3.2 459 2 US-09-097-889-22 Sequence 59, Appl
36 88 3.2 549 2 US-08-676-279-59 Sequence 11, Appl
37 87.5 3.2 471 4 US-09-032-742-11 Sequence 14, Appl
38 87.5 3.2 471 4 US-09-032-742-14 Sequence 4, Appl
39 87.5 3.2 707 2 US-08-576-165-4 Sequence 2, Appl
40 87.5 3.2 764 1 US-08-424-567-2 Sequence 2, Appl
41 87.5 3.2 764 4 US-08-711-928-2 Sequence 2, Appl
42 87.5 3.2 764 4 US-09-184-937-2 Sequence 58, Appl
43 87 3.2 576 2 US-08-676-279-58 Sequence 17, Appl
44 86.5 3.1 471 4 US-09-032-742-17 Sequence 2, Appl
45 86 3.1 672 1 US-07-841-651-2

ALIGNMENTS

RESULT 1
US-08-677-734A-12
; Sequence 12, Application US/08677734A
; Patent No. 5871919
; GENERAL INFORMATION:
; APPLICANT: Brant, Steven R.
; APPLICANT: Yun, Chris C.H.
; APPLICANT: Donowitz, Mark
; APPLICANT: Tse, Chung-Ming
; TITLE OF INVENTION: Cloning, Tissue Distribution, and
; TITLE OF INVENTION: Functional Analysis Of The Human Na+/H+ Exchanger Isoform,
; TITLE OF INVENTION: NHE3.
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,734A
; FILING DATE: 10-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0043-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 832 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-677-734A-12

Query Match 17.1%; Score 471; DB 2; Length 832;
Best Local Similarity 29.1%; Pred. No. 4.8e-39;
Matches 148; Conservative 96; Mismatches 197; Indels 68; Gaps 19;

		TELEPHONE: (202) 408-4000	
		TELEFAX: (202) 408-4400	
		INFORMATION FOR SEQ ID NO: 11:	
		SEQUENCE CHARACTERISTICS:	
		LENGTH: 831 amino acids	
		TYPE: amino acid	
		STRANDEDNESS: single	
		TOPOLOGY: linear	
		MOLECULE TYPE: peptide	
		US-08-677-734A-11	
		Query Match 17.0%; Score 467; DB 2; Length 831;	
		Best Local Similarity 28.3%; Pred. No. 1.2e-38;	
		Matches 144; Conservative 99; Mismatches 205; Indels 60; Gaps 18;	
QY	20	VVALNLFVALLCACIV-LGHLLLENRMNESITALLIGLGTGTGVTILLISKGSSHLVLF-	77
Db	56	IIALWLVASLAKIVFHLSH--KVTSVVPESALLIVGLVGGIVL-----AADHIAST	108
QY	78	-SEDLFFIYLLPPIIFNAGFOVKKKOFRRNFVTIMLFGAVGTIISCTIISLGVTOFFKKL	136
Db	109	LPTVTFEFLPPIVIDAGYFMPNLFESNLGSLILLYAVGTVMNAATTGLSLGVFSLG	168
QY	137	DIGTFDLG--DYLAIGAIFAATDSVCTQLVNO-DETPLLYSLVFGGVNDATSVVVEN	193
Db	169	INGELKIGLLDPLFGSLIAAOPVAVLAVFEHVHNEVLFIIVFGESLLNDVAVTVLYN	228
QY	194	ATQSFDTLHNEAAPHLLGNFLYFLLSLLGAATGLISAYVIKKL-YFGRHSTDREVA	252
Db	229	VQSFVTLLGGDVTGDCVKGISSFVVS--LGGTLGVGVFAFLLSLVTRFTKRVRIE	287
QY	253	LMMLAYLSYMLAEFLDLSGILTVFCGIVMSHYTHNVNTESSRIITKHTFATLSFLAET	312
Db	288	FVFLISYLSLSEMLSSLSLAILATFCGICCKYKANISEOSATTVRYTMKMLASGAET	347
QY	313	FIFYVGMDALDIDKRSVSDTPGTSIAVSSILMGLVMVGRAAFVFPFLSFLSNLAKKNOS	372
Db	348	IIFMELGISAVDPLIW-----TWNTAFVRLTLL--FVSFRAIGVVQLTGLLNRYRMVQL	400
QY	373	EKINFNMQVVIWWSGLMRGAVSMALAY---NKFTAGTIDVRGNAIMTSTITVCLFST	428
Db	401	ELID-----QVVMYSGGL-RGAVAFVALLDGNK-----VKENLNFVSTIIIVVFTV	448
QY	429	VVFGMLTKPLISYLLPHONATTSMLSDNTPK---SIHIPLDQDSFIEPSSGNHNVPRPD	485
Db	449	IFOGLIKPLVQWLKVKR-----SEHREPKLNEKLHGRAFD-----HILSAIEDISG	491
QY	486	SIRGFTTRTRTVHYV-----WRQFDDSEFM	510
Db	492	DISG-----QIGHNYLRDKWANEFRRL	514
RESULT 2			
US-08-677-734A-11			
; Sequence 11, Application US/08677734A			
; Patent No. 5871919			
; GENERAL INFORMATION:			
; APPLICANT: Brant, Steven R.			
; APPLICANT: Yun, Chris C.H.			
; APPLICANT: Donowitz, Mark			
; APPLICANT: Tse, Chung-Ming			
; TITLE OF INVENTION: Cloning, Tissue Distribution, and			
; TITLE OF INVENTION: Functional Analysis of The Human Na+/H+ Exchanger Isoform,			
; TITLE OF INVENTION: NHE3.			
; NUMBER OF SEQUENCES: 12			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &			
; ADDRESSEE: Dunner			
; STREET: 1300 I Street, N.W., Suite 700			
; CITY: Washington			
; STATE: D.C.			
; COUNTRY: USA			
; ZIP: 20005-3315			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/677,734A			
; FILING DATE: 10-JUL-1996			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Fordis, Jean B.			
; REGISTRATION NUMBER: 32,984			
; REFERENCE/DOCKET NUMBER: 05387.0043-00000			
; TELECOMMUNICATION INFORMATION:			

ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677,734A
FILING DATE: 10-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0043-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-677-734A-9

Query Match 16.8%; Score 462; DB 2; Length 834;
Best Local Similarity 28.6%; Pred. No. 4e-38;
Matches 146; Conservative 97; Mismatches 203; Indels 64; Gaps 18;

Qy 20 VVALNLFVALLCACIVLGHLEENRMNESITALLIGLGTGVTILLISKG---KSSHLV 76
Db 57 VIALWILVASLAK---IGFHLSHKVTSSVVPESALLIVLG-----LVLGGIWVAADHTAS 107
Qy 77 F--SEDLFFIYLLPPIIFNAGFOVKKQFFRNFVITMLFGAVGTIISCTIISLGVTOFFK 134
Db 108 FTLTPVTFVFFLLPPIVLDAGYFMPNRLFGNLGTLIYAVVGTVMNAATTGLSLYGVFL 167
Qy 135 KLDIGTFDLG--DYLAIGAIFAATDSVCTLOVLNO--DETPLYLSLVFGGVNDATSVVV 191
Db 168 SGLMGDLQIGLDFLLFGSLMAADVPVAVLAVFEVHVNEVLFIIVFGESLLNDVTVVL 227
Qy 192 FNAIQSFDTLHLNHEAFLHLLGNFLYLLSTLLGAATGLISAYVKKL-YFGRHSTDR 250
Db 228 YNPFESVALGGDNVTGDCVKGIIVFFVVS-LGSTLVGVVFAFLSLVTRFKHVRIE 286
Qy 251 VALMMLMAYLSYMLAEFLDGLTFFCGIVMNSHYTHNVTSSRITTKKTEATLSFLA 310
Db 287 PGFVFIISYLSYLTSEMLSALAITFCGICCKYKVNANISEQATVRYTMKMLASSA 346
Qy 311 ETFLFYVGMADLIDKWRVSDTPGTSTAVSSILMGLVMGAAVFPFLSFLSLAKKN 370
Db 347 ETIIFMFLGISAVNPFIW-----TWNTAFVLLTV--FISVYRAIGVVLQTLWLLNRYRMV 399
Qy 371 QSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTAGHTDVRGNAMITSTITVCLFSTVV 430
Db 400 QLEPID---QVLSYGGCL-RGAVAFALV---VLLDGDVKVKEKNLFVSTTIIVVFTVIF 451
Qy 431 FGMUTKPLISYLLPHONATTSMISDNTPK---SIHIPLLDQDSITEPSGNHNVRPDSI 487
Db 452 QGLTKPLVQWLKVR-----SEHREPLNEKLHGRAFD-----HILSAIEDI 494
Qy 488 RGLTRTPRTRVHYV---WROFDDSMRPV 513
Db 495 SG-----QIGHNVLKDKWSHFDRKFLSRV 518

RESULT 4

US-08-677-734A-10
; Sequence 10, Application US/08677734A
; Patent No. 5871919
; GENERAL INFORMATION:

APPLICANT: Brant, Steven R.
APPLICANT: Yun, Chris C.H.
APPLICANT: Donowitz, Mark
APPLICANT: Tse, Chung-Ming
TITLE OF INVENTION: Cloning, Tissue Distribution, and
TITLE OF INVENTION: Functional Analysis Of The Human Na+/H+ Exchanger Isoform,
TITLE OF INVENTION: NHE3.
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677,734A
FILING DATE: 10-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0043-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-677-734A-10

Query Match 16.8%; Score 462; DB 2; Length 834;
Best Local Similarity 28.6%; Pred. No. 4e-38;
Matches 146; Conservative 97; Mismatches 203; Indels 64; Gaps 18;

Qy 20 VVALNLFVALLCACIVLGHLEENRMNESITALLIGLGTGVTILLISKG---KSSHLV 76
Db 57 VIALWILVASLAK---IGFHLSHKVTSSVVPESALLIVLG-----LVLGGIWVAADHTAS 107
Qy 77 F--SEDLFFIYLLPPIIFNAGFOVKKQFFRNFVITMLFGAVGTIISCTIISLGVTOFFK 134
Db 108 FTLTPVTFVFFLLPPIVLDAGYFMPNRLFGNLGTLIYAVVGTVMNAATTGLSLYGVFL 167
Qy 135 KLDIGTFDLG--DYLAIGAIFAATDSVCTLOVLNO--DETPLYLSLVFGGVNDATSVVV 191
Db 168 SGLMGDLQIGLDFLLFGSLMAADVPVAVLAVFEVHVNEVLFIIVFGESLLNDVTVVL 227
Qy 192 FNAIQSFDTLHLNHEAFLHLLGNFLYLLSTLLGAATGLISAYVKKL-YFGRHSTDR 250
Db 228 YNPFESVALGGDNVTGDCVKGIIVFFVVS-LGSTLVGVVFAFLSLVTRFKHVRIE 286
Qy 251 VALMMLMAYLSYMLAEFLDGLTFFCGIVMNSHYTHNVTSSRITTKKTEATLSFLA 310
Db 287 PGFVFIISYLSYLTSEMLSALAITFCGICCKYKVNANISEQATVRYTMKMLASSA 346
Qy 311 ETFLFYVGMADLIDKWRVSDTPGTSTAVSSILMGLVMGAAVFPFLSFLSLAKKN 370
Db 347 ETIIFMFLGISAVNPFIW-----TWNTAFVLLTV--FISVYRAIGVVLQTLWLLNRYRMV 399
Qy 371 QSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTAGHTDVRGNAMITSTITVCLFSTVV 430
Db 400 QLEPID---QVLSYGGCL-RGAVAFALV---VLLDGDVKVKEKNLFVSTTIIVVFTVIF 451

[illegible]

```

1 RESULT
2
3 US-09-238-796-2
4 ; Sequence 2, Application US/09238796
5 ; Patent No. 6074845
6 ;
7 GENERAL INFORMATION:
8 ;
9 APPLICANT: AIYAR, NAMBI V.
10 ;
11 APPLICANT: DISA, JYOTI
12 ;
13 TITLE OF INVENTION: RECLR·BOVINE CALCITONIN RECEPTOR-LIKE
14 ;
15 TITLE OF INVENTION: RECEPTOR
16 ;
17 NUMBER OF SEQUENCES: 2
18 ;
19 CORRESPONDENCE ADDRESS:
20 ;
21 ADDRESSEE: Ratner & Prestia
22 ;
23 STREET: P.O. Box 980
24 ;
25 CITY: Valley Forge
26 ;
27 STATE: PA
28 ;
29 COUNTRY: USA
30 ;
31 ZIP: 19482
32 ;
33 COMPUTER READABLE FORM:
34 ;
35 MEDIUM TYPE: Diskette
36 ;
37 COMPUTER: IBM Compatible
38 ;
39 OPERATING SYSTEM: DOS
40 ;
41 SOFTWARE: Fastseq for Windows Version 2.0
42 ;
43 CURRENT APPLICATION DATA:
44 ;
45 APPLICATION NUMBER: US/09/238,796
46 ;
47 FILING DATE: 28-JAN-1999
48 ;
49 CLASSIFICATION:
50 ;
51 PRIOR APPLICATION DATA:
52 ;
53 APPLICATION NUMBER:
54 ;
55 FILING DATE:
56 ;
57 ATTORNEY/AGENT INFORMATION:
58 ;
59 NAME: Prestia, Paul F
60 ;
61 REGISTRATION NUMBER: 23,031
62 ;

```

; REFERENCE/DOCKET NUMBER: GP-70599
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0700
; TELEFAX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-238-796-2

Query Match 3.6%; Score 99; DB 3; Length 462;
Best Local Similarity 20.9%; Pred. No. 0.12;
Matches 90; Conservative 59; Mismatches 130; Indels 152; Gaps 23;

QY 13 STSDHASV-VALLFVALLCAGVIGLHLEENRWNESITALLIGLGTGVTILLISCKGS 71
Db : : | | | | | : : : : : | | | | | : : : : :
129 NVNTHKEKVTALNFY-----LTIIGHVL-----SIASLLISLG----IFFYFKSL 171
QY 72 SHLVFSEDLFFIYLLPPIIFNAGVQKKOFFERFVIMLFGAVG---TIISCTIISLG 128
Db : : | | | | | : : : : : | | | | | : : : : :
172 CORITHLKNLFVFCNSVI-----TIHLTAVANNQALVATNPVSCK 214
QY 129 VTQFFKKLDIGTDLGDLAIGAIFAATDSVCTLOVQNODETPLLISLVFGGV-VNDAT 187
Db : : | | | | | : : : : : | | | | | : : : : :
215 VSQFIHLXLMG-----CN-----YFNMCEGIVLHFLV 242
QY 188 SVVFNIAQSFDLTHLNEAAPHLLGNFLYFLLSLILCAATGLISA---YVINKLFGR 244
Db : : | | | | | : : : : : | | | | | : : : : :
243 VVAVFAEKO-----HLM---WYF-----LGMGFPLIPACIHAVARRLYND 281
QY 245 H---STDREVALMMLMAYLSYLAELFDLSGILTVFFCGIVMSHYTHNVTESSRIITKH 301
Db : : | | | | | : : : : : | | | | | : : : : :
282 NCWISSDTQLLYIHGPICALLVNLFFLLNIVRVLITKLKVTQAESNLYMKA---VRA 338
QY 302 TFATLSFLAETFLVGMGDAIDDKRSVSDTPTGTSIA-----VSSILM---GLVMVG 352
Db : : | | | | | : : : : : | | | | | : : : : :
339 TLILVPLLGIEFVLI-----PWR-----PEGKAEIYDIYIIMLMHYOGLLV-- 381
QY 353 RAAVFPPLSFLSLAKKNQSE-KINFNMQVVIWISGLRGAVSMALAYNKETFRAGHTDVR 411
Db : : | | | | | : : : : : | | | | | : : : : :
382 STIFCFEFGVEQAILRRNNQYKIQFG-----NNFS---HSDTL 417
QY 412 GNAMITSRIT 422
Db : : | | | : : : : :
418 RSASYTVSTIS 428

RESULT 7
US-09-141-2
; Sequence 2, Application US/08098141
; Patent No. 5441875
; GENERAL INFORMATION:
; APPLICANT: Hediger Ph.D., Matthias A.
; TITLE OF INVENTION: Urea Transporter Polypeptide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:

; APPLICATION NUMBER: US/08/098.141
; FILING DATE: 19930723
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan Esq., Warren A.
; REGISTRATION NUMBER: 34,199
; REFERENCE/DOCKET NUMBER: 092662-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-227-5020
; TELEFAX: 617-227-7566
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 397 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-098-141-2

Query Match 3.6%; Score 98.5; DB 1; Length 397;
Best Local Similarity 21.3%; Pred. No. 0.11;
Matches 68; Conservative 49; Mismatches 73; Indels 129; Gaps 17;

QY 34 IVLGHLEENRWNESITALLIGLGTGVTILLISCKGS-----H-----LLVFS 78
Db 80 IVIG-LFVNQPMW--AIAGCLGTVMSTTALLISQDRASIALSGHGYNGVLVGLLIAVFS 136
QY 79 E--DLFFIYLLPPIIFNAGVQKKOFFERFVIMLFGAVGTIISCTIISLVGTQFFKKL 136
Db : : | | | | | : : : : : | | | | | : : : : :
137 DKGDYVWLLLPVIMS-----MSCPLSSALGTIFSKW 170
QY 137 DIGTFDLGDLAIGAIFAATDSVCTLOVQNODETPLLISLVFGGVVNDATSV--VVFNA 194
Db : : | | | | | : : : : : | | | | | : : : : :
171 DPVFTLPENIATVLYLAATGH-----YNLFFPTTLQPVSSVNPNTWSE 215
QY 195 IQSFDLTHLNEAAPHLL-----GNFLY-LFLLSLT-----LGA 227
Db : : | | | | | : : : : : | | | | | : : : : :
216 IQ-----VPLLRLRAIPVGIGVYGCNDPWTGGIFLIALFISSPLICLHAAIGS 263
QY 228 ATCLISAYVI---KKLYEGRHSTDREVA-----LMLMAYLSYMLA---ELF----- 268
Db 264 TGMGLAALTIATPDSIYFGLCGFNSTLACIAGVGMFYVITWQTHLLAVACALFAAYVGA 323
QY 269 DLSGILTVE-----FC 279
Db : : | | | : : : : :
324 ALTNVLSVFGPLTCTWPEC 342

RESULT 8
US-08-472-028A-4
; Sequence 4, Application US/08472028A
; Patent No. 5767373
; GENERAL INFORMATION:
; APPLICANT: Ward, Eric R
; APPLICANT: Volrath, Sandra
; TITLE OF INVENTION: Manipulation of Protoporphyrinogen
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.028A
; FILING DATE:

```

; APPLICATION NUMBER:  US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIORITY DATE: 28-FEB-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:  US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:  US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME:  Meigs, J. Timothy
; REFERENCE/DOCKET NUMBER:  38,241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  (919) 541-8587
; TELEFAX:  (919) 541-8689
; INFORMATION FOR SEQ ID NO:  4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 508 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-808-931-4
;
; Query Match 3.4%; Score 95; DB 2; Length 508;
; Best Local Similarity 18.5%; Pred. No. 0.36;
; Matches 55; Conservative 48; Mismatches 87; Indels 108; Gaps
;
; QY 290 NVTSSRITTKHFAFPLSLFAETFIPLYG----MDALDIDKWSVSDTPTGTSIAVSSLL 345
; DB 181 SAADPDSLSMKHSPDILWNVEKSGFGLIIVGARTTFAAKGKSRDTKSSFGTKKG---- 235
; QY 346 MGLVMVGGRAAFVFP--LSFL-SNLAKKNQSEKINFNMQVVIWISGLMRGAVSMALAYNKF 402
; DB 236 -----SRGSEFSGGMQILPDLTCLCKSLSHDEINLDSKV-----LSLSYNSG 276
; QY 403 TR-----AGHTDVRG-----NAIMTSTITVC----- 424
; DB 277 SRQENWLSVCVSHNETQRQNPHYDAVIMTAPL--CNVEMKVMKGGQFPQLNLPLEINYM 334
; QY 425 ----LFTSVWFSGMITKPLISY--LLPHQONAT-----TSMLSDDNTPKSTHPLLD 468
; DB 335 PLSVLITITFTKEKYKRPLEGFGLVLPFSKEQKGFCTGLTFSSMFPDRSPSDVHL----- 390
; QY 469 QDSFTEPGSNVINPR--PDSIRGELT-----RTRTVHYVWRQ----FDDSF 509
; DB 391 YTTFIGGRNOELAKASTDELKQVTVSDIQRLLGVEGEPVSVNHYHWRKAPPLYDSSY 448
;
; RESULT 10
; US-08-808-323-4
; Sequence 4, Application US/08808323
; Patent No. 6018105
; GENERAL INFORMATION:
; APPLICANT: Johnson, Marie
; APPLICANT: Volrath, Sandra
; APPLICANT: Ward, Eric
; TITLE OF INVENTION: Promoters from Plant
; TITLE OF INVENTION: Protoporphyrogen Oxidase Genes
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 6018105artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/808,323
FILING DATE: 28-FEB-1996
CLASSIFICATION: 800
PRIOR APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 508 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-323-4

Query Match 3.4%; Score 95; DB 3; Length 508;
Best Local Similarity 18.5%; Pred. No. 0.36;
Matches 55; Conservative 48; Mismatches 87; Indels 108; Gaps 15;

QY 290 NVTSSRTTKHTPATLSFLAETFIPLYG----MDALDIDKWRVSVDTPGTSIAVSSIL 345
DB 181 SAADPDSLKMKHSPDUNNVEKSGSIIVGAIKTKAAGKGRDRTKSPGKKG----- 235
QY 346 MGLVMVGRAAFVFP--LSFL-SNLAKKNQSEKINFNQVVMWGLMRGAVSMALAYNKF 402
DB 236 -----SRGSEFSGKGMQILPDLTKLSLHDEINLDSKV-----LSLSYNSG 276
QY 403 TR-----AGHTDVRG-----NAIMITITIVC----- 424
DB 277 SRQENWLSVCVSHNETQRPHYDAVIMTAPL--CNVKEMKVMKGQPFQNLPLPEINYM 334
QY 425 ----LFSTVVFGLTKPLISY--LLPHQNTAT-----TSMLSDDNTPKSIHIPLLD 468
DB 335 PLSVLITFTTEKVKRPLEGFGVLPSKEQKHGFKTLGTLFSSMMFPDRSPSDVHL---- 390
QY 469 QDSFIEPSGNHNVP--PDSIRGFLT-----RPRTRVHYWRO-----FDDSF 509
DB 391 YTTFIGGSRNOELAKASTDELKQVWTSDLQRLILGVEGEPVSVNHYWYRKAFPLYDSSY 448

RESULT 11

US-09-050-603A-4
Sequence 4, Application US/09050603A
Patent No. 6023012

GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Heifetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
TITLE OF INVENTION: Protoporphyrogen Oxidase
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6023012artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,603A
FILING DATE: 30-MAR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 508 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-050-603A-4

Query Match 3.4%; Score 95; DB 3; Length 508;
Best Local Similarity 18.5%; Pred. No. 0.36;
Matches 55; Conservative 48; Mismatches 87; Indels 108; Gaps 15;

QY 290 NVTSSRTTKHTPATLSFLAETFIPLYG----MDALDIDKWRVSVDTPGTSIAVSSIL 345
DB 181 SAADPDSLKMKHSPDUNNVEKSGSIIVGAIKTKAAGKGRDRTKSPGKKG----- 235
QY 346 MGLVMVGRAAFVFP--LSFL-SNLAKKNQSEKINFNQVVMWGLMRGAVSMALAYNKF 402
DB 236 -----SRGSEFSGKGMQILPDLTKLSLHDEINLDSKV-----LSLSYNSG 276
QY 403 TR-----AGHTDVRG-----NAIMITITIVC----- 424
DB 277 SRQENWLSVCVSHNETQRPHYDAVIMTAPL--CNVKEMKVMKGQPFQNLPLPEINYM 334
QY 425 ----LFSTVVFGLTKPLISY--LLPHQNTAT-----TSMLSDDNTPKSIHIPLLD 468
DB 335 PLSVLITFTTEKVKRPLEGFGVLPSKEQKHGFKTLGTLFSSMMFPDRSPSDVHL---- 390
QY 469 QDSFIEPSGNHNVP--PDSIRGFLT-----RPRTRVHYWRO-----FDDSF 509
DB 391 YTTFIGGSRNOELAKASTDELKQVWTSDLQRLILGVEGEPVSVNHYWYRKAFPLYDSSY 448

RESULT 12

US-09-102-4208-4
Sequence 4, Application US/091024208
Patent No. 6084155

GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Ward, Eric
APPLICANT: Heifetz, Peter
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
TITLE OF INVENTION: OXIDASE ("PROTOX")


```

Db      391 YTTFIGGSRNQELAKASTDELKQVVTSDLQRLLGVGEPEPSVNNHYWRKAFFLYDSSY   448

RESULT      15
US-09-015-683-4
; Sequence 4, Application US/09015683
; Patent No 6288306
; GENERAL INFORMATION:
; APPLICANT: Ward, Eric R
; APPLICANT: Volrath, Sandra
; TITLE OF INVENTION: Manipulation of Protoporphyrinogen
; TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,683
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1748/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 508 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-015-683-4

Query Match          3.4%; Score 95; DB 4; Length 508;
Best Local Similarity 18.5%; Pred. No. 0.36;
Matches 55; Conservative 48; Mismatches 87; Indels 108; Gaps 15

Qy      290 NVYTESRITTKTTFATLSLATFTFLYVG-----MDALDDIKWRSVDTFCTSIAYSSIL 345
       : : : | | : : : | : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      181 SAADPDSLKMKSFPDLNHWVEKSGSIIVGAIRTKFAAGCKSDTKSPGTKKG----- 235
       : : : | | : : : | : : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy      346 MGLVMVGRAAFVFP--LSPL-SNLAKKNQSEKINFMQVVIWSGLMRGAVSMALAYNKF 402
       : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      236 -----SRGSFEFGQMQLPTDLCKLSLHDEINLDKV-----LSSLVNSG 276
       : : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy      403 TR-----AGHTDVRG-----NAIMITSTITVC----- 424
       : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db      277 SRQENNSLCVSCHNETQRONPHYDAVIMTAPL--CNVKEMKVMKGQPQLNFLPEINYM 334
       : : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |

Qy      425 ---LSTVVVFGMLTKPLISY--LLPHONAT-----TSLMSDDNTPKSTHIPLLD 468
       : | : : | : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      335 PLSVLIITTTKEKVRPLGGFVLIPSREQKHGFKTGLTFSMMFDRGPSDVHL----- 390
       : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |

Qy      469 QDSFIPEGSGNNHVP--PDSIRGFLT-----RPRTVHHYYWRQ----FDOSF 509
       : | | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db      391 YTTFIGGSRNQELAKASTDELKQVVTSDLQRLLGVGEPEPSVNNHYWRKAFFLYDSSY   448

Search completed: November 17, 2001, 13:27:59
Job time: 82 sec

```

	QY	469 QDSFTIEPSGNHNVP--PDISIRGFLT-----RPTRTVHYWRO----FDDSF 509 :: : : : :
	Db	391 YTTFIGGSNOELAKASTDELKQQVTSDLORLLGVGEPEVSNNHYYWRKAFLPYDSSY 448 :
	RESULT	14
	US-09-196-268-4	:
	; Sequence 4, Application US/09196268	:
	; Patent No. 6282837	:
	; GENERAL INFORMATION:	:
	; APPLICANT: Ward, Eric R	:
	; APPLICANT: Volrath, Sandra	:
	; TITLE OF INVENTION: Manipulation of Protoporphyrinogen	:
	; TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms	:
	; NUMBER OF SEQUENCES: 12	:
	; CORRESPONDENCE ADDRESS:	:
	; ADDRESSEE: Ciba-Geigy Corporation	:
	; STREET: 7 Skyline Drive	:
	; CITY: Hawthorne	:
	; STATE: NY	:
	; COUNTRY: USA	:
	; ZIP: 10532	:
	; COMPUTER READABLE FORM:	:
	; MEDIUM TYPE: Floppy disk	:
	; COMPUTER: IBM PC compatible	:
	; OPERATING SYSTEM: PC-DOS/MS-DOS	:
	; SOFTWARE: Patentin Release #1.0, Version #1.25	:
	; CURRENT APPLICATION DATA:	:
	; APPLICATION NUMBER: US/09/196,268	:
	; FILING DATE: 06-JUN-1995	:
	; CLASSIFICATION:	:
	; PRIOR APPLICATION DATA:	:
	; APPLICATION NUMBER: US 08/261,198	:
	; FILING DATE: 16-JUN-94	:
	; ATTORNEY/AGENT INFORMATION:	:
	; NAME: Elmer, James Scott	:
	; REGISTRATION NUMBER: 36,129	:
	; REFERENCE/DOCKET NUMBER: CGC 1748/CIP	:
	; TELECOMMUNICATION INFORMATION:	:
	; TELEPHONE: 919-541-8614	:
	; TELEFAX: 919-541-8689	:
	; INFORMATION FOR SEQ ID NO: 4:	:
	; SEQUENCE CHARACTERISTICS:	:
	; LENGTH: 508 amino acids	:
	; TYPE: amino acid	:
	; TOPOLOGY: linear	:
	; MOLECULE TYPE: protein	:
	US-09-196-268-4	:
	Query Match 3.4%; Score 95; DB 4; Length 508;	:
	Best Local Similarity 18.5%; Pred. Nismatches 87; Indels 108; Gaps	:
	Matches 55; Conservative 48; Mismatches 87; Indels 108; Gaps	:
	QY 290 NTVESSRIITKHFFATILSFLAETFIPLYVG---MDALDIKWRSVDTGTCTSIIVSSL 345 ; : : : : : :	
	Db 181 SAADPDLSLUMSKHSPPDLWNNEKSFGSIIGAITRTFAAKGGKRSDTKSPCTKKG --- 235 ; : : : : : :	
	QY 346 MGLVMVGRGAAPFP--LSFL-SMLAKKNSEKINFNMQVIWWSGLMRGAVSMALAYNF 402 ; : : : : : :	
	Db 236 -----SRGFSPFKGMQLLPDTLKSLSHDEINLDKV-----LSLSYNSG 276 ; : : : : : :	
	QY 403 TR-----AGHTDVKG-----NAMITSTTVTC----- 424 ; : : : : : :	
	Db 277 SRQENWSLCVSHNETQRQNPHFYDAVIMTAPL--CNVKEMKMVKGGQPFOLFNLPEINTM 334 ; : : : : : :	
	QY 425 ----LFSTVVPGMLTKPLISY--LLPHQNAI-----TMLSDNDTPKSIHIPILED 468 ; : : : : : :	
	Db 335 PLSVLIITTFTEKVKRPLEGFGVLIPSKEOKHGPKTLTGTFSSMMFPDRSPDHV---- 390 ; : : : : : :	
	QY 469 QDSFTIEPSGNHNVP--PDISIRGFLT-----RPTRTVHYWRO----FDDSF 509 :: : : : :	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2001, 13:26:57 ; Search time 27.58 Seconds
(without alignments)
2580.859 Million cell updates/sec

Title: US-09-271-584A-2
Perfect score: 2755
Sequence: 1 MDSLVSKLSTSLSDHASV.....FVPFVPGSPTRNPPDLKA 538

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_16:*

1: sp.archaea:*

2: sp.bacteria:*

3: sp.fungi:*

4: sp.human:*

5: sp.invertebrate:*

6: sp.mammal:*

7: sp.mhc:*

8: sp.organelle:*

9: sp.phage:*

10: sp.plant:*

11: sp.rodent:*

12: sp.unclassified:*

13: sp.invertebrate:*

14: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2755	100.0	538	10	Q9ZPK3	Q9ZPK3 arabidopsis
2	2168.5	78.7	555	10	Q9F2N0	Q9F2N0 atp1a1a
3	2136	77.5	457	10	Q04655	Q04655 arabidopsis
4	2041	74.1	542	10	Q9F8E3	Q9F8E3 pharbitis n
5	2023.5	73.4	535	10	Q9SXX8	Q9SXX8 oryza sativ
6	1910.5	69.3	552	10	Q9SQ00	Q9SQ00 arabidopsis
7	1567	56.9	529	10	Q9FJ63	Q9FJ63 arabidopsis
8	615.5	22.3	561	5	Q9XZ44	Q9XZ44 drosophila
9	615.5	22.3	649	5	Q9VPE1	Q9VPE1 drosophila
10	586	21.3	140	10	Q9LKH5	Q9LKH5 mesembryant
11	583.5	21.2	703	5	Q9XW14	Q9XW14 caenorhabdi
12	576.5	20.9	616	4	Q75827	Q75827 homo sapien
13	570	20.7	727	5	Q9WM99	Q9WM99 drosophila
14	568	20.6	687	5	Q9U624	Q9U624 drosophila
15	544	19.7	569	3	Q13726	Q13726 schizosacch
16	542	19.7	153	10	Q9LKH6	Q9LKH6 mesembryant
17	531	19.3	629	5	Q20944	Q20944 caenorhabdi
18	494.5	17.9	437	4	Q9V507	Q9V507 homo sapien
19	489.5	17.0	560	3	Q9HEX3	Q9HEX3 pneumocysti

20	458.5	16.6	518	5	Q9VIF9	Q9VIF9 drosophila
21	451	16.4	411	4	Q9Y2E8	Q9Y2E8 homo sapien
22	451	16.4	1203	5	Q9NGZ4	Q9NGZ4 drosophila
23	442.5	16.1	1179	5	Q9NCQ0	Q9NCQ0 aedes aegypt
24	432	15.7	673	5	Q23706	Q23706 carcinus ma
25	424	15.4	478	5	Q01995	Q01995 caenorhabdi
26	414	15.0	698	13	Q9W724	Q9W724 cyprinus ca
27	414	15.0	813	13	Q9W6Q1	Q9W6Q1 amphiuma tr
28	411.5	14.9	779	13	Q9W714	Q9W714 platichthys
29	410	14.9	781	13	P70009	P70009 xenopus lae
30	406	14.7	798	5	Q21386	Q21386 caenorhabdi
31	391.5	14.2	634	5	Q9T2H0	Q9T2H0 caenorhabdi
32	386.5	14.0	602	5	Q09432	Q09432 caenorhabdi
33	385.5	14.0	651	5	Q16452	Q16452 caenorhabdi
34	366.5	13.3	660	5	Q23617	Q23617 caenorhabdi
35	315.5	11.5	768	13	Q9W7S0	Q9W7S0 anguilla an
36	311.5	11.3	494	1	Q29412	Q29412 archaeoglob
37	289	10.5	684	5	Q19444	Q19444 caenorhabdi
38	279.5	10.1	531	10	Q9WA14	Q9WA14 arabidopsis
39	278.5	10.1	527	2	P73863	P73863 synechocyst
40	269	9.8	524	2	Q9LCB5	Q9LCB5 bacillus su
41	259	9.4	530	2	Q9F3L8	Q9F3L8 streptomyce
42	258	9.4	1146	10	Q9LKW9	Q9LKW9 arabidopsis
43	254.5	9.2	575	10	Q9SLJ7	Q9SLJ7 arabidopsis
44	244	8.9	581	2	Q9HXX9	Q9HXX9 pseudomonas
45	241	8.7	581	2	Q9Z9H1	Q9Z9H1 pseudomonas

ALIGNMENTS

RESULT	ID	PRELIMINARY;	PRT;	538 AA.
Q9ZPK3	AC	Q9ZPK3		
Q9ZPK3	DT	01-MAY-1999 (Tremblrel. 10, Created)		
Q9ZPK3	DT	01-MAY-1999 (Tremblrel. 10, Last sequence update)		
Q9ZPK3	DT	01-MAY-2000 (Tremblrel. 13, Last annotation update)		
Q9ZPK3	DE	SODIUM PROTON EXCHANGER NHX1 (FRAGMENT).		
Q9ZPK3	GN	NHE1.		
Q9ZPK3	OS	Arabidopsis thaliana (Mouse-ear cross).		
Q9ZPK3	OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
Q9ZPK3	OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;		
Q9ZPK3	OC	Brassicales; Brassicaceae; Arabidopsi.		
Q9ZPK3	OX	NCBI_TaxID=3702;		
Q9ZPK3	RN	[1]		
Q9ZPK3	RP	SEQUENCE FROM N.A.		
Q9ZPK3	RC	STRAIN=CV. COLUMBIA;		
Q9ZPK3	RX	MEDLINE=99145575; PubMed=9990049;		
Q9ZPK3	RA	Gaxiola R.A., Rao R., Sherman A., Grisafi P., Alper S.L., Fink G.R.;		
Q9ZPK3	RT	"The Arabidopsis thaliana proton transporters, AtNHX1 and AtPL1, can		
Q9ZPK3	RL	function in cation detoxification in yeast."		
Q9ZPK3	RL	Proc. Natl. Acad. Sci. U.S.A. 96:1480-1485(1999).		
Q9ZPK3	RN	[2]		
Q9ZPK3	RP	SEQUENCE FROM N.A.		
Q9ZPK3	RA	Quintero F.J., Blatt M.R., Pardo J.M.;		
Q9ZPK3	RT	"The AtNHX1 gene encodes a putative Na+/H+ antiporter."		
Q9ZPK3	RL	Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.		
Q9ZPK3	DR	EMBL; AF106324; AAD16946.1; -		
Q9ZPK3	DR	EMBL; AF096190; AAF21755.1; -		
Q9ZPK3	DR	InterPro; IPR0000576; -		
Q9ZPK3	DR	InterPro; IPR001179; -		
Q9ZPK3	DR	Pfam; PF00999; Na_H_Exchange; 1.		
Q9ZPK3	DR	PRINTS; PR01084; NAHEXCHNGR.		
Q9ZPK3	DR	PROSITE; PS00453; FKBP_PP1ASE_1; UNKNOWN_1.		
Q9ZPK3	FT	NON_TER		
Q9ZPK3	SQ	SEQUENCE 538 AA; 59513 MW; 1189AD6C5C726996 CRC64;		

Query Match 100.0%; Score 2755; DB 10; Length 538;
Best Local Similarity 100.0%; Pred. No. 3.7e-169;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MLDLSVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 60
DB 1 MLDLSVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 60
QY 61 VTILLISGKSSHLLVSEDLFFIYLLPPIIFNAGFQVKKOFFRNFVIMLFGAVGTII 120
DB 61 VTILLISGKSSHLLVSEDLFFIYLLPPIIFNAGFQVKKOFFRNFVIMLFGAVGTII 120
QY 121 SCTIISLGVTQFFKKLDIGTDLGDLAIGAIFAATDSVCTQLQVNDQETPLLSLVFGE 180
DB 121 SCTIISLGVTQFFKKLDIGTDLGDLAIGAIFAATDSVCTQLQVNDQETPLLSLVFGE 180
QY 181 GVVNDATSVVFNIAQSFDTLHNLHNEAAPHLLGNFLYLLSTLLGAATGLISAVYIKKL 240
DB 181 GVVNDATSVVFNIAQSFDTLHNLHNEAAPHLLGNFLYLLSTLLGAATGLISAVYIKKL 240
QY 241 YFGRHSTDRVALMMLMAYLSYMLAEFLDLSGLTTFVFCGIVMWSHYTHWNTVTESSRIITK 300
DB 241 YFGRHSTDRVALMMLMAYLSYMLAEFLDLSGLTTFVFCGIVMWSHYTHWNTVTESSRIITK 300
QY 301 HTFAPLSLAETFTFLYVGMADLIDKWRVSVDTPGTSIAVSSILMGLVMGAAAFVPEPL 360
DB 301 HTFAPLSLAETFTFLYVGMADLIDKWRVSVDTPGTSIAVSSILMGLVMGAAAFVPEPL 360
QY 361 SFLSNLAKKNSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTTRAGHTDVRGNAMITST 420
DB 361 SFLSNLAKKNSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTTRAGHTDVRGNAMITST 420
QY 421 ITVCLFSTVTFVGMILTKPLISYLLPHQNAATT--SMLSDNTPKSIHIPILD--QDSFIE-P 475
DB 421 ITVCLFSTVTFVGMILTKPLISYLLPHQNAATT--SMLSDNTPKSIHIPILD--QDSFIE-P 475
QY 481 VPRPDSIRGFLTRPTRVVHYWQFDDSEMRPVFGGGRGVFVPGSPTRNPPDLSCA 538
DB 481 VPRPDSIRGFLTRPTRVVHYWQFDDSEMRPVFGGGRGVFVPGSPTRNPPDLSCA 538

```

RESULT 2

```

QY Q9FZNO PRELIMINARY; PRT; 555 AA.
AC Q9FZNO;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE NA/H ANTI-PORTER NHX1.
GN AGNHX1.
OS Atriplex gmelini.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
OC Caryophyllales; Chenopodiaceae; Atriplex.
OX NCBI_TaxID=118074;
RN [1]
RP SEQUENCE FROM N.A.
RA Hamada A., Hayakawa T., Shono M., Xia T., Hayashi Y., Tanaka A.;
RT "Isolation and characterization of a Na/H antiporter gene from the
RT halophyte Atriplex gmelini."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038492; BAB11940.1;
DR EMBL: 555 AA; 61504 MW; ACD5ED45FF3D398A CRC64;
SQ SEQUENCE

```

Query Match 78.7%; Score 2168.5; DB 10; Length 555;
 Best Local Similarity 77.7%; Pred. No. 1.6e-131;
 Matches 425; Conservative 53; Mismatches 56; Indels 13; Gaps 5;

```

QY 2 LDSLVS-KLPSLSTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 60
DB 5 LSSLGKMDALTTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 64
QY 61 VTILLISGKSSHLLVSEDLFFIYLLPPIIFNAGFQVKKOFFRNFVIMLFGAVGTII 120
DB 65 VVILLISGKSSHLLVSEDLFFIYLLPPIIFNAGFQVKKOFFRNFVIMLFGAVGTIV 124

```

```

QY 121 SCTIISLGVTQFFKKLDIGTDLGDLAIGAIFAATDSVCTQLQVNDQETPLLSLVFGE 180
DB 125 SFTIISLGAISIFKKLDIGTDLGDLAIGAIFAATDSVCTQLQVNDQETPLLSLVFGE 184
QY 181 GVVNDATSVVFNIAQSFDTLHNLHNEAAPHLLGNFLYLLSTLLGAATGLISAVYIKKL 240
DB 185 GVVNDATSVVFNIAQSFDTLHNLHNEAAPHLLGNFLYLLSTLLGAATGLISAVYIKKL 244
QY 241 YFGRHSTDRVALMMLMAYLSYMLAEFLDLSGLTTFVFCGIVMWSHYTHWNTVTESSRIITK 300
DB 245 YFGRHSTDRVALMMLMAYLSYMLAEFLDLSGLTTFVFCGIVMWSHYTHWNTVTESSRIITK 304
QY 301 HTFAPLSLAETFTFLYVGMADLIDKWRVSVDTPGTSIAVSSILMGLVMGAAAFVPEPL 360
DB 305 HAFAPLSLAEVAFVFLYVGMADLIDKWRVSVDTPGTSIAVSSILMGLVMGAAAFVPEPL 364
QY 361 SFLSNLAKKNSEKINFNMQVVIWWSGLMRGAVSMALAYNKFTTRAGHTDVRGNAMITST 420
DB 365 SWLMNFAKKSOSEKVTNQOQIVIWWSGLMRGAVSMALAYNQFTRSGHTQLRCNAIMITST 424
QY 421 ITVCLFSTVTFVGMILTKPLISYLLPHQNAATT--SMLSDNTPKSIHIPILD--QDSFIE-P 475
DB 425 ISVLFSTVTFVGMILTKPLISYLLPHQNAATT--SMLSDNTPKSIHIPILD--QDSFIE-P 484
QY 476 SGNH-----NVPRPDSIRGFLTRPTRVVHYWQFDDSEMRPVFGGGRGVFVPGSPTRNPPDLSCA 538
DB 485 NGNHEDTTEPTIVRPSIRMLLNAPHTTVHYWQFDDSEMRPVFGGGRGVFVPGSPTRNPPDLSCA 544
QY 529 ERNPPDL 535
DB 545 EOSTNNL 551

```

RESULT 3

```

QY 004655 PRELIMINARY; PRT; 457 AA.
AC 004655;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE SIMILAR TO SODIUM/HYDROGEN EXCHANGER.
GN A.TW021B04.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Dante M., Wansley P., Gibson A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF007271; AAB61069.1;
DR InterPro; IPR000676;
DR InterPro; IPR001179;
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PROSITE; PS00453; FKBP_PPIASE_1; UNKNOWN_1.
SQ SEQUENCE 457 AA; 50611 MW; 0AF2F235F1A259EE CRC64;

```

Query Match 77.5%; Score 2136; DB 10; Length 457;
 Best Local Similarity 95.6%; Pred. No. 1.5e-129;
 Matches 430; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

```

QY 1 MLDLSVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 60
DB 1 MLDLSVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLEENRWNNESITALLIGLGTG 60

```

QY 61 VTILLISGKSHLLVSEDLFFIYLLPPIIFNAGFOVKKOFFRNFVTIMLFGAVGTII 120
 DB 61 VTILLISGKSHLLVSEDLFFIYLLPPIIFNAGFOVKKOFFRNFVTIMLFGAVGTII 120
 QY 121 SCTLISGLVTOFFKKLDIGTFDGLDYLAIGAIPAATDSVCTLOVLNODETPLLYSLVFE 180
 DB 121 SCTLISGLVTOFFKKLDIGTFDGLDYLAIGAIPAATDSVCTLOVLNODETPLLYSLVFE 180
 QY 181 GVVNDATSVVVFNAIQSFDTLHNLHNEAARHLLGNFLYLLFLLSTLLGAA----- 228
 DB 181 GVVNDATSVVVFNAIQSFDTLHNLHNEAARHLLGNFLYLLFLLSTLLGAAVPLFSSLPFEL 240
 QY 229 TGLISAYVKKLYFG-----RHSDREVALMMLAYLSYMLAELELDSGLTVFCG 280
 DB 241 TGLISAYVKKLYFGPHINCHRHSDREVALMMLAYLSYMLAELELDSGLTVFCG 300
 QY 281 IVMSHYTHWNTYESSRITTKHTATISFLAETFIYLVGMALDIDKWRVSVDTPGTSIA 340
 DB 301 IVMSHYTHWNTYESSRITTKHTATISFLAETFIYLVGMALDIDKWRVSVDTPGTSIA 360
 QY 341 VSSILMGLVMVGRAAFVPLSFLSNLAKKNOSEKINFNMVVIWWSGLMRGAVSMALAYN 400
 DB 361 VSSILMGLVMVGRAAFVPLSFLSNLAKKNOSEKINFNMVVIWWSGLMRGAVSMALAYN 420
 QY 401 KTRAGHTDVRGNAMITSTIITVCLFSTVV 430
 DB 421 KTRAGHTDVRGNAMITSTIITVCLFSTVV 450

RESULT 4
 QYFEB3 PRELIMINARY; PRT; 542 AA.
 AC QYFEB3:
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE NA+/H+ EXCHANGER.
 GN PURPLE.
 OS Pharbitis nil (Violet) (Japanese morning glory).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Convolvulaceae; Ipomoea.
 OX NCBI_TaxID=35883;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=KZSK-2; TISSUE=LEAF, AND FULLY COLORED FLOWER BUDS;
 RX MEDLINE=20487008; PubMed=11034195;
 RA Fukada-Tanaka S., Inagaki Y., Yamauchi T., Saito N., Iida S.;
 RT "Colour-enhancing protein in blue petals.";
 RL Nature 407:581-581(2000).
 DR EMBL; AB033980; BAB16381.1; -;
 DR EMBL; AB033989; BAB16380.1; -;
 SQ SEQUENCE 542 AA; 59973 MW; 4B47FDE04401A191 CRC64;

Query Match 74.1%; Score 2041; DB 10; Length 542;
 Best Local Similarity 73.8%; Pred. No. 2.3e-123;
 Matches 399; Conservative 57; Mismatches 71; Indels 14; Gaps 4;

QY 2 LDSLVSKLPSLSTSDHASVVALNLFVALLCACIVLGHLEENRWNESITALLIGLCTGV 61
 DB 5 LSSLQN-SDLFTSDHASVSNVSNLFCACIVLGHLEENRWNESITALLIGLCTGV 63
 QY 62 TILLISGKSHLLVSEDLFFIYLLPPIIFNAGFOVKKOFFRNFVTIMLFGAVGTIIS 121
 DB 64 VILLISGKSHLLVSEDLFFIYLLPPIIFNAGFOVKKOFFRNFVTIMLFGAVGTIIS 123
 QY 122 CTISISGVTOFFKKLDIGTFDGLDYLAIGAIPAATDSVCTLOVLNODETPLLYSLVFE 181
 DB 124 CSISGAVKIFKHLIDIDFDGLDYLAIGAIPAATDSVCTLOVLNODETPLLYSLVFE 183
 QY 182 VVNDATSVVVFNAIQSFDTLHNLHNEAARHLLGNFLYLLFLLSTLLGAAVPLFSSLPFEL 241
 DB 182 VVNDATSVVVFNAIQSFDTLHNLHNEAARHLLGNFLYLLFLLSTLLGAAVPLFSSLPFEL 241

DB 184 VVNDATSVVVFNAIQSFDMTSFDPKIGLHIGNFLYLLFLLSTLLGAAVPLFSSLPFEL 243
 QY 242 FGRHSTDRVALMMLAYLSYMLAELELDSGLTVFCGIVMSHYTHWNTYESSRITTKH 301
 DB 244 FGRHSTDRVALMMLAYLSYMAELFVLSGLTVFCGIVMSHYTHWNTYESSRITTKH 303
 QY 302 TFATLSLAETFIYLVGMALDIDKWRVSVDTPGTSIAVSSILMGLVMVGRAAFVPLS 361
 DB 304 SFATLSVAETFIYLVGMALDIDKWRVSVDTPGTSIAVSSILMGLVMVGRAAFVPLS 363
 QY 362 FLSNLAKKNOSEKINFNMVVIWWSGLMRGAVSMALAYNKETFRAGHTDVRGNAMITSTI 421
 DB 364 FLSNLAKKNOSEKINFNMVVIWWSGLMRGAVSMALAYNKETFRAGHTDVRGNAMITSTI 423
 QY 422 TVCLFSTVVFGMLTKPLSYLLP-----HONATTSMLSDDNTPKSIHPLLDQ--DS 471
 DB 424 TVCLFSTVVFGMLTKPLSYLLP-----HONATTSMLSDDNTPKSIHPLLDQ--DS 480
 QY 472 FIEPSGNHNPDPDSIRGFTRTVHYWQFDSFMRPVFGGRGVFPVFGSPPTERN 531
 DB 481 ESDMITGPEVARTALRMILLRTPTHTVHYWQFDSFMRPVFGGRGVFPVFGSPPTERN 540
 QY 532 P 532
 DB 541 P 541

Query Match 73.4%; Score 2023.5; DB 10; Length 535;
 Best Local Similarity 73.3%; Pred. No. 3e-122;
 Matches 389; Conservative 60; Mismatches 79; Indels 3; Gaps 2;

QY 2 LDSLVSKLPSL-STSDHASVVALNLFVALLCACIVLGHLEENRWNESITALLIGLCTG 60
 DB 3 MEVAAARIGALYTTSDVASVSNLFCACIVLGHLEENRWNESITALLIGLCTG 62
 QY 61 VTILLISGKSHLLVSEDLFFIYLLPPIIFNAGFOVKKOFFRNFVTIMLFGAVGTII 120
 DB 63 VVILLMTKGSHLVSEDLFFIYLLPPIIFNAGFOVKKOFFRNFVTIMLFGAVGTMI 122
 QY 121 SCTLISGVTOFFKKLDIGTFDGLDYLAIGAIPAATDSVCTLOVLNODETPLLYSLVFE 180
 DB 123 SFTTISIAAIFRMDITDVGDFLAIGAIFSAIDSVCTLOVLNODETPLLYSLVFE 182
 QY 181 GVVNDATSVVVFNAIQSFDTLHNLHNEAARHLLGNFLYLLFLLSTLLGAAVPLFSSLPFEL 240

```

Db 183 GYVNDATSIIVFNALONFDLHDAAVLVPLGNFFFLSSTFLGVFAGLLSAYIIKKL 242
QY 241 YFGRHSTDEVALMMLMAYLSYMLAEFLDLSGILTVFPCGIVMASHYTHWNTVTESSRITTK 300
Db 243 YGRHSTDEVALMMLMAYLSYMLAEFLDLSGILTVFPCGIVMASHYTHWNTVTESSRITTK 302
QY 301 HFATLSFLAEFIPLYVGMALDIDKRSVSDTPGTSIAVSSILMGLVMVGRAAFVPL 360
Db 303 HFATLSFLAEFIPLYVGMALDIDKRSVSDTPGTSIAVSSILMGLVMVGRAAFVPL 362
QY 361 SFLSNLAKKNOSEKINFNQVVMWGLMRGAVSMALAYNKFTFRAGHTDVRGNAMITST 420
Db 363 SFLSNLTKKAPKEKITRQVQVVMWGLMRGAVSMALAYNKFTFRAGHTDVRGNAMITST 422
QY 421 ITVCLFSTVVFGLMFKPLISILYLP--HONATFSM--LSDDNTPKSIHIPILLDDQDSTFEPS 476
Db 423 ITVCLFSTVVFGLMFKPLISILYLP--HONATFSM--LSDDNTPKSIHIPILLDDQDSTFEPS 476
QY 481 VPRPDSIRGLFTRPTRTVHYVWQDSDFMRPVFGGFGFVFPVPGSPTERN 531
Db 481 IVPSSLRMLLKPTHTVHYVWQDSDFMRPVFGGFGFVFPVPGSPTERN 531

RESULT 6
Q9SQU0 PRELIMINARY; PRT; 552 AA.
AC Q9SQU0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE PUTATIVE SODIUM PROTON EXCHANGER.
GN F24P17.16.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Konning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011623; AAF08577.1; -.
DR InterPro; IPR000676; -.
DR InterPro; IPR001179; -.
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PROSITE; PS00453; FKBP_LPTASE_1; UNKNOWN_1.
SQ SEQUENCE 552 AA; 61135 MW; FB4317D8A874FCE9 CRC64;

Query Match 69.3%; Score 1910.5; DB 10; Length 552;
Best Local Similarity 69.7%; Pred. No. 5.5e-115;
Matches 377; Conservative 58; Mismatches 97; Indels 9; Gaps 3;

QY 2 LDSLVKSLSTSDHASVVALNLFVALLCACIVLGHLLLEENRWNNESITALLIGLGTGV 61
Db 5 LSTMLEKTEALFASDHASVSNLNFVALLCACIVLGHLLLEENRWNNESITALLIGSGTGI 64
QY 62 TILLISKGKSSHLVSEDLFFIYLLPIFNAGFQVKKQFFRNVFTMLFGAVGTIIS 121
Db 65 VILLISGGKSSRLVSEDLFFIYLLPIFNAGFQVKKQFFRNVFTMLFGAVGTIIS 124
QY 122 CTIISLGVQTFKKLIDGTFDGLYLAIGATFAATSDVCTQLQVNDDETPLLYSLVFVGG 181
Db 125 FVLIISFGAKHLFEKNWIGDITADYLAIGATFAATSDVCTQLQVNDDETPLLYSLVFVGG 184
QY 182 VVNDATSVVFNATIOSFDLTHLNEAFHLLGNFLYLLFLLSTLLGATGLISAYVYKKLY 241

```

```

Db 185 VVNDATSVVFNATIOSFDLTHLNEAFHLLGNFLYLLFLLSTLLGATGLISAYVYKKLY 244
QY 242 FGRHSTDEVALMMLMAYLSYMLAEFLDLSGILTVFPCGIVMASHYTHWNTVTESSRITTK 301
Db 245 IGRHSTDEVALMMLLAYSVMYLAELFHLSSILTVFPCGIVMASHYTHWNTVTESSRITTK 304
QY 302 TPATLSFLAEFIPLYVGMALDIDKRSVSDTPGTSIAVSSILMGLVMVGRAAFVPL 361
Db 305 TPATLSFLAEFIPLYVGMALDIDKRSVSDTPGTSIAVSSILMGLVMVGRAAFVPL 364
QY 362 FUSNLAKKNOSEKINFNQVVMWGLMRGAVSMALAYNKFTFRAGHTDVRGNAMITST 421
Db 365 FUSNLTKSPDEKIDLKQVTVIWMGLMRGAVSMALAYNQFTTSHTKVLGNAMITST 424
QY 422 TVCLFSTVVFGLMFKPLISILYLP--HONATFSM--LSDDNTPKSIHIPILLDDQDSTFEPS 476
Db 425 TVCLFSTVVFGLMFKPLISILYLP--HONATFSM--LSDDNTPKSIHIPILLDDQDSTFEPS 476
QY 477 GNNVPRPDSIRGLFTRPTRTVHYVWQDSDFMRPVFGGFGFVFPVPGSPTERNPPDLS 536
Db 485 PRQHYV---SFRMEFKSPSRAIHVYWRKFDNAYMRIFGGRGVSPVFGSPVFGSPVQWS 540
QY 537 K 537
Db 541 E 541

RESULT 7
Q9FJ63 PRELIMINARY; PRT; 529 AA.
AC Q9FJ63;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE SODIUM PROTON EXCHANGER.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=99087489; PubMed=9872454;
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
RT Sequence features of the regions of 1,013,767 bp covered by sixteen
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:297-308(1998).
DR EMBL; AB015479; BAB08564.1; -.
SQ SEQUENCE 529 AA; 58871 MW; 8A6BD42E822F90C1 CRC64;

Query Match 56.9%; Score 1567; DB 10; Length 529;
Best Local Similarity 58.3%; Pred. No. 5.6e-93;
Matches 307; Conservative 80; Mismatches 120; Indels 20; Gaps 5;

QY 15 SDHASVVALNLFVALLCACIVLGHLLLEENRWNNESITALLIGLGTGVILLISKGKSSHL 74
Db 15 ACHPOQIVISFVAILCLVIGHLLLEENRWNNESITALLIGVAAAGSTVILLISKGKSSHI 74
QY 75 LVFSEDLFFIYLLPIFNAGFQVKKQFFRNVFTMLFGAVGTIISLGVTFQFFK 134
Db 75 LVFDEELFFIYLLPIFNAGFQVKKQFFRNVFTMLFGAVGTIISLGVTFQFFK 134
QY 135 KLIDGTFDGLYLAIGATFAATSDVCTQLQVNDDETPLLYSLVFVGGVNDATSVVFN 194
Db 135 KLGFGLSARDYLAIGATFAATSDVCTQLQVNDDETPLLYSLVFVGGVNDATSVVFN 194
QY 195 IQSFDLTHLNEAFHLLGNFLYLLFLLSTLLGATGLISAYVYKKLYFCRHSSTOREVALM 254
Db 195 VQKIQFESLGTALQVFGNGLYLLFLLSTLLGATGLISAYVYKKLYFCRHSSTOREVALM 254

```

QY 255 MLAYLSYMAELFDLSGLTVFFCGIWMYSHYTHNVTSSRITTKHTFATLSLAEFTFI 314
 DB 255 VLAYLSYMAELFDLSGLTVFFCGIWMYSHYTHNVTSSRITTKHTFATLSLAEFTFI 314
 QY 315 FLYVGMALDIDKRSVSDTPGTSITAVSSILMGLVVMVGRAAAFVPLSFLNLAKN--QS 372
 DB 315 FLYVGMALDIDKRSVSDTPGTSITAVSSILMGLVVMVGRAAAFVPLSFLNLAKN--QS 372
 QY 315 FLYVGMALDIDKRSVSDTPGTSITAVSSILMGLVVMVGRAAAFVPLSFLNLAKN--QS 372
 DB 315 FLYVGMALDIDKRSVSDTPGTSITAVSSILMGLVVMVGRAAAFVPLSFLNLAKN--QS 372
 QY 373 EKINFMQVVIWMSGLMGAVSMALAYNKFTFRAGHTDVRGNAMITITVCLFSTVVFV 432
 DB 373 EKINFMQVVIWMSGLMGAVSMALAYNKFTFRAGHTDVRGNAMITITVCLFSTVVFV 432
 QY 375 ESITPKHQVVIWMSGLMGAVSMALAYNKFTFRAGHTDVRGNAMITITVCLFSTVVFV 434
 DB 375 ESITPKHQVVIWMSGLMGAVSMALAYNKFTFRAGHTDVRGNAMITITVCLFSTVVFV 434
 QY 433 MLTKPLISVLLP-----HQNATSMLSDDNTPK-SIHPLLDQDSEFSPGNHNVPRDPSI 487
 DB 433 MLTKPLISVLLP-----HQNATSMLSDDNTPK-SIHPLLDQDSEFSPGNHNVPRDPSI 487
 QY 435 FUTKPLVNLVLLPQDASHNTGNGRKTGPGSKEDATLPL-----SFDESASTNFRAKDSI 491
 DB 435 FUTKPLVNLVLLPQDASHNTGNGRKTGPGSKEDATLPL-----SFDESASTNFRAKDSI 491
 QY 488 RGLFTRPTTHVYWRQDDGFMRFVFGVGRFVPGSPPTERNPPD 534
 DB 488 RGLFTRPTTHVYWRQDDGFMRFVFGVGRFVPGSPPTERNPPD 534
 QY 492 SLLMEQPVVTHRYWRKEDDYMRFPG-----PRENOPE 528
 DB 492 SLLMEQPVVTHRYWRKEDDYMRFPG-----PRENOPE 528
 RESULT 8
 Q9XZ4 PRELIMINARY: PRT; 561 AA.
 ID Q9XZ4
 AC Q9XZ4
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-NAR-2001 (Tremblrel. 16, Last annotation update)
 DE SODIUM-HYDROGEN EXCHANGER NHE1.
 GN NHE1 OR CG12178.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dow J.A.T.;
 RT "An insect member of the Na⁺/H⁺ exchanger family";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF142676; AAD32689.1;
 DR FlyBase: FBgn026787; Nhe1.
 DR InterPro: IPR000676;
 DR Pfam: PF00999; Na_H_Exchange; 1.
 DR PRINTS: PR01084; NAHEXCHNGR.
 DR SEQUENCE 561 AA; 60974 MW; 6B53FC7A92528627 CRC64;
 QY 10 PSLTSDHASV-----ALNLFVALLCACIVLHLEEN 43
 DB 63 PSISAGNASTTKRGNASTLVTPDPLDISHAVEGHNSSLSLF-VTCV-IMLGILLHS 120
 QY 44 -----RWNNESITALLIGLGTGVITLLIS--KGKSSHLVFESEDLFFYLLPPIFNAG 95
 DB 121 MLQTGFQYLPESIVVFLGAFGLSLNVMGONGSKWKEEVFSPMGFLVLLPPIFESG 180
 QY 96 FOVKKKQFRRNVTTMLGAVGTIISCTIISLGTQVFFKLDIG-----TDLGDYLAIGA 151
 DB 181 YNLHKGNFQNTGSLVFAIFGTTTISALVIGAGI-----YLLGLGEVAFRLSFSSESFAFGS 236
 QY 152 IFAATDSVCTLOVNO-DTEPLLYSLVFGVGVNDATSVVVFNAI-OSFDLTHLNHEA- 208
 DB 237 LISADVPATVAIFHALVDPLNMLVFGESILNDASIVLTASITOS-----ANNVNAEAST 293
 QY 209 ----PHLLGNFLYFLLLSTLLGAATGLISAYVKKLYFGHRSTDEVALMMLAYLSYML 264
 DB 294 GEAMFSALKTCFCAFFASAGIGVIFALISALLKHLIDLRKHS-LEFAMLMFTYAPYVL 352

QY 265 AELFDLSGLTVFFCGIWMYSHYTHNVTSSRITTKHTFATLSLAEFTFI 324
 DB 353 ABGILGIMAILFCGIMYSHYTHNVTSSRITTKHTFATLSLAEFTFI 412
 QY 325 IDKWSVSDDTPGTSITAVSSILMGLV--MVGRAAFVPLSFLNLAKNKSEKINENMOVV 382
 DB 413 FKH-----QVELSFVIAIVLCIGRACNIFPLAFLVN---KFRCHKNNKMOFI 459
 QY 383 IWSGLMGAVSMALAYNKFTFRAGHTDVRGNAMITITVCLFSTVVFVGLTKPLI 439
 DB 460 MWFSGL-RGASIALSL-----HLNLDSEKRRHVITITLIIVLFTLVLGSTMPLL 511
 QY 440 SYLLP 444
 DB 512 KYLKP 516
 RESULT 9
 Q9VPJ1 PRELIMINARY: PRT; 649 AA.
 ID Q9VPJ1
 AC Q9VPJ1
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE NHE1 PROTEIN
 GN NHE1 OR CG12178.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Lilang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Plattman G.S., Pan S., Poillard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).

```
DR EMBL; AE003590; AAF51559.1; -.
DR FlyBase; FBgn0026787; NheI.
DR InterPro; IPR000676; -.
DR Pfam; PF00999; Na.H.Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
SQ SEQUENCE 649 AA; 71282 MW; 920499A2D8C27ADD CRC64;

Query Match 22.3%; Score 615.5; DB 5; Length 649;
Best Local Similarity 33.6%; Pred. No. 7.5e-32;
Matches 163; Conservative 93; Mismatches 148; Indels 81; Gaps 18;

Qy 10 PSLTSDHSAW-----ALNFVALLCACIVLGHLEEN 43
Db 63 PSISASGNASTKRGNASTLVTOPPLIDSHAVEQEHNSLSLFF-VICV-IMLGILLIHS 120
Qy 44 -----RWNESITALLIGLGTGVTILLIS--KGRSSHLVFESEDLFFIYLLPPIIFNAG 95
Db 121 MLQTCFVLPESIVVFLGAFIGLSLNVMSQNGSKREEVSPMGFFLVLLPPIIFESG 180
Qy 96 FOVKKOFRNFVIMLFGAVGIIISCTIISLGVTOFFKKLDIG----TFDLGDLAIGA 151
Db 181 YNLHKGNFQIGSILVFATFGTISALVIGAGI-----YLLGLGEVAFRLSFSFAPGS 236
Qy 152 IFAATDSVCTLOVLNQ--DETPLYLSLVFGEVVDATSVVVENAI-QSFDLTLNHNHAAA- 208
Db 237 LISAVDPATVAIPHALDVPILNMLVFGESILNDAISIVLTASITQS---ANVNAEAST 293
Qy 209 ----PHLLGNFLYFLSLTLGGAATGLISAYVIKLYFGHSTDRVLAEMLMAYLSYML 264
Db 294 GEAMFSALKTCAMFFASAGIVFALISALLKHIDLRKHPES-LEFAMLMFTYAPYVL 352
Qy 265 AELFDLSGLTVFCGVIMSHYTHWNVTSSRTTKHTFATLSFLAEFIYLYVGMALD 324
Db 353 AEGHLSGIMALEFCGVIMSHYTHNLSVTQITMQOTMRTLAFAETCVFAYLGLAIFS 412
Qy 325 IDKRSVSDPTGTSIAVSSILMGV--WVGRAAFVPLSFLSNLAKKQSEKINFNQV 382
Db 413 FKH-----QVELSFVIMAVICLGRACNIPFLAFLVN---KPREHKINKKMOFI 459
Qy 383 IWSGLMRGAVSMALAYNKETFRAGTHDVRGNA---IMTSTITVCLFSTVVFVGMLTPLI 439
Db 460 MWFSGSL-RGALSVALSL-----HLNLDSEKRRHVIITTLIIIVFTLVLGSGTWPLL 511
Qy 440 SYLKP 444
Db 512 KYLKP 516

RESULT 10
Q9LKH5 PRELIMINARY; PRT; 140 AA.
AC Q9LKH5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PUTATIVE Na+/H+-ANTIORTER (FRAGMENT).
GN NHX2.
OS Mesembryanthemum crystallinum (Common ice plant).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC Caryophyllales; Alzooceae; Mesembryanthemum.
OX NCBI_TaxID=3544;
RN [1]
RP SEQUENCE FROM N.A.
RA Chauhan S., Forsthoefel N., Bohnert H.J.;
RT "Na+/myo-inositol symporters and Na+/H+-antiporters in plant sodium
RT transport.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF279671; AAF91424.1; -.
DR InterPro; IPR001179; -.
DR PROSITE; PS00453; FKBP_PP1ASE_1; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 703 AA; 76942 MW; 9C054503CEED767 CRC64;

Query Match 21.2%; Score 583.5; DB 5; Length 703;
Best Local Similarity 34.6%; Pred. No. 9.3e-30;
Matches 150; Conservative 79; Mismatches 160; Indels 45; Gaps 12;

Qy 25 LFVALLCACIVLGHLEENRWNESITALLIGLGTGVTILLISKGRSSHLVFESEDLFFI 84
Db 138 LFVIMLATLVVHMLIVSKIIHMPESAIVAGLALIG-SILSYSRDWDSEIALSFDVFL 196
Qy 85 YLLPPIIFNAGFOVKKOFRNFVIMLFGAVGIIISCTIISLGVTOFFKKLDIGITFDLG 144
Db 197 VLLPPIIFENAYNLNKGFFSNFVILTFATGTTISAMVIGAGLYILGAILGIFETFF 256
FT NON_TER 1
SQ SEQUENCE 703 AA; 76942 MW; 9C054503CEED767 CRC64;

Query Match 21.3%; Score 586; DB 10; Length 140;
Best Local Similarity 80.0%; Pred. No. 1e-30;
Matches 112; Conservative 16; Mismatches 12; Indels 0; Gaps 0;

Qy 100 KQOFRNFVIMLFGAVGIIISCTIISLGVTOFFKKLDIGITFDLGDLAIGAIFAATDSV 159
Db 1 KQOFRNFVIMLFGAVGIIISCTIISLGVTOFFKKLDIGITFDLGDLAIGAIFAATDSV 60
Qy 160 CTLOVLNODETPLLYSLVFGVVDATSVVVENAIQSFDLTLNHNHAFHLLGNFLYLF 219
Db 61 CTLOVLNODETPLLYSLVFGVVDATSVVVENAIQSFDLTLNHNHAFHLLGNFLYLF 120
Qy 220 LLSTLLGAATGLISAYVIKK 239
Db 121 FTSTLLGAITGLSAYIIRK 140

RESULT 11
Q9XW14 PRELIMINARY; PRT; 703 AA.
AC Q9XW14;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Y18D10A.6 PROTEIN.
GN Y18D10A.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38 (1994).
DR EMBL; AL034393; CAA22320.1; -.
DR InterPro; IPR000676; -.
DR Pfam; PF00999; Na.H.Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
SQ SEQUENCE 703 AA; 76942 MW; 9C054503CEED767 CRC64;
```



```
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003614; AAF52423.2; -.
DR FlyBase; FBgn0028703; Nhe3.
DR InterPro; IPR000676; -.
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
SQ SEQUENCE 727 AA; 80331 MW; 36D730CC7C3B17B CRC64;

Query Match 20.7%; Score 570; DB 5; Length 727;
Best Local Similarity 29.3%; Pred. No. 7.1e-29;
Matches 178; Conservative 109; Mismatches 169; Indels 132; Gaps 24;

QY 9 LPSLSTSDHASVVALNL--FVALLCACIVLGHLLLENR--WNNESTITALLIGLGTGVTIL 124
Db 65 LQAKATLNH-RIQSLDLLVFEVLLALTTLTWLFKHHRVSWLHETGLAVIYGLVGAIR 123
QY 58 --GTGVITLLIS-----KGRSSHLLV----- 76
Db 124 YAGTSATLVHMQVEPQGVPTYSKLPDPDTLWFRYPVNTNGTKPPEGIKTYAYVFRGOVH 183
QY 77 -----FSEDLFFIYLLPIIFNAGFOVKKKOFFRNFVTIMLFGAVGTIISCTI 124
Db 184 DVDENEIDLKATDPPEVFNILPIIFVAGYSLKKYFFRNGLGAILTFAIVGTTLISAF 243
QY 125 IS---LGVTOFFKLDIGTDLGDLAIGAIFAATDSVCTLOVLNODETPL-LYSLVFG 180
Db 244 IGGFMVGCVKLMPKYLSSSFTFDSLYFGALISPTDPLTLAIFNDLRVDVNLVYALV 303
QY 181 GVNDATSVVFNFAIQSFDTLHNLHAAFHL-----LGNFLYLLFLLSTLLGAATGLISA 234
Db 304 SVLNDVAIVLSGAIQNG--EHSNTGEFETAPLRSLSDFFSIFLLSLMIGAAMGLTA 362
QY 235 YVIKKLYFGRHSTDREVALMMLMAYLSYMLAELFDSGLTVPFGIVMSHYTHWNVTES 294
Db 363 -LMTKTRVRDPFLESALFVLSYFTFLAEATELTGVVAVLFCGICQAHYIYNLSD 421
QY 295 SRTTKHTFATLSFLAETFFIFYVGMALDIDKWRVSVDPTGTSIAVSSILMGLV--MVG 352
Db 422 SRQTKQIFELLNLAENFIYSIGV-----SMFTPKHHFDAGFIITAFICAAIG 472
QY 353 RAAVFVPLSPSLNAKKNOSEKINFNMQVIVWWSGLMRGAVSMALAYNKFTTRAGHTDVRG 412
Db 473 RAVNVYPLSWLLNKKR---PKISTNFQHMLFFAGL-RGAMSFALA-----IRNTVSDER- 523
QY 413 NAIMTITIVCLFSTVVFVGMILTKPLISYL-LPHONATTSMLSDDDNTPKSIHIPLLDQDS 471
Db 524 -QTMLTATSLIVFTVVIQGAANFLNLKIP-----VGVDDTEQLNNYQVHSSDG 575
QY 472 FIEP-----SG--NHNVPRP-DSIRGFLTRPT-----RTVH-----YYWRQF 505
Db 576 YLQDVEGGVNRNKMRLSGGDTNLDTPVDGTNGSLGASGGRRRNSHEKAILARIWGNF 635
QY 506 DDSFMRPV 513
Db 636 DTKYMKPL 643

RESULT 14
ID Q9U624 PRELIMINARY; PRT; 687 AA.
AC Q9U624;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE SODIUM-HYDROGEN EXCHANGER NHE3.
GN NHE3 OR CC11328.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
```

```

[1]
RN SEQUENCE FROM N.A.
RP Giannakou M.E., Dow J.A.T.;
RT "A novel Drosophila member of the NHE exchanger family.";
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF199463; AAF13702.1; -.
DR FlyBase; FBgn0028703; Nhe3.
DR InterPro; IPR000676; -.
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
SQ SEQUENCE 687 AA; 75775 MW; C5D6BA9BDDCBA1DE CRC64;

Query Match 20.6%; Score 568; DB 5; Length 687;
Best Local Similarity 29.3%; Pred. No. 8.9e-29;
Matches 178; Conservative 105; Mismatches 193; Indels 132; Gaps 23;

QY 9 LPSLSTSDHASVVALNL--FVALLCACIVLGHLLLENR--WNNESTITALLIGLGTGVTIL 64
Db 25 LQAKATLNH-RIQSLDLLVFEVLLALTTLTWLFKHHRVSWLHETGLAVIYGLVGAIR 83
QY 65 LI-SKGRSSHLLV----- 76
Db 84 YAGTSGLDHHMQVEPQGVPTYSKLPDPDTLWFRYPVNTNGTKPPEGIKTYAYVFRGOVH 143
QY 77 -----FSEDLFFIYLLPIIFNAGFOVKKKOFFRNFVTIMLFGAVGTIISCTI 124
Db 144 DVDENEIDLKATDPPEVFNILPIIFVAGYSLKKYFFRNGLGAILTFAIVGTTLISAF 203
QY 125 IS---LGVTOFFKLDIGTDLGDLAIGAIFAATDSVCTLOVLNODETPL-LYSLVFG 180
Db 204 IGGFMVGCVKLMPKYLSSSFTFDSLYFGALISPTDPLTLAIFNDLRVDVNLVYALV 263
QY 181 GVNDATSVVFNFAIQSFDTLHNLHAAFHL-----LGNFLYLLFLLSTLLGAATGLISA 234
Db 264 SVLNDVAIVLSGAIQNG--EHSNTGEFETAPLRSLSDFFSIFLLSLMIGAAMGLTA 322
QY 235 YVIKKLYFGRHSTDREVALMMLMAYLSYMLAELFDSGLTVPFGIVMSHYTHWNVTES 294
Db 323 -LMTKTRVRDPFLESALFVLSYFTFLAEATELTGVVAVLFCGICQAHYIYNLSD 381
QY 295 SRTTKHTFATLSFLAETFFIFYVGMALDIDKWRVSVDPTGTSIAVSSILMGLV--MVG 352
Db 382 SRQTKQIFELLNLAENFIYSIGV-----SMFTPKHHFDAGFIITAFICAAIG 432
QY 353 RAAVFVPLSPSLNAKKNOSEKINFNMQVIVWWSGLMRGAVSMALAYNKFTTRAGHTDVRG 412
Db 433 RAVNVYPLSWLLNKKR---PKISTNFQHMLFFAGL-RGAMSFALA-----IRNTVSDER- 483
QY 413 NAIMTITIVCLFSTVVFVGMILTKPLISYL-LPHONATTSMLSDDDNTPKSIHIPLLDQDS 471
Db 484 -QTMLTATSLIVFTVVIQGAANFLNLKIP-----VGVDDTEQLNNYQVHSSDG 535
QY 472 FIEP-----SG--NHNVPRP-DSIRGFLTRPT-----RTVH-----YYWRQF 505
Db 536 YLQDVEGGVNRNKMRLSGGDTNLDTPVDGTNGSLGASGGRRRNSHEKAILARIWGNF 595
QY 506 DDSFMRPV 513
Db 596 DTKYMKPL 603

RESULT 15
ID O13726 PRELIMINARY; PRT; 569 AA.
AC O13726;
DT 01-NOV-1998 (TremBLrel. 08, Created)
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TremBLrel. 13, Last annotation update)
DE PUTATIVE SODIUM/HYDROGEN EXCHANGER C15A10.06.
GN SPAC15A10.06.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
```


Search completed: November 17, 2001, 13:29:00
Job time: 123 sec

source 1. .2313
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 297. .1904
 /gene="OsNHX1"
 297. .1904
 /note="similar to yeast sodium/proton exchanger"
 /codon_start=1
 /protein_id="BA83337.1"
 /db_xref="GI:5731737"
 /translation="MGMEVAAARLALYTTSDYASVVSINLVALLCACVILGHLEEE
 NKNWNSIALIIGLCTGVILLMTKSKSHLFSEDFIYLLPPIIPNAGQVKK
 KOFFNFMTILFGAVGTMISEFFTSIAAIAFRSMNIGTLDVDFLAIGAFESATDS
 VCTLOVMODTFLYSLVFGEGVNDATLIVLENALQNFDLVHDAVYVFKLGNFF
 YLFLSFLGVAGLLSAYIIKKLYIGRHSFDREVALMLMLVLSYMLATLIDLSCIL
 TVFFGIVMSWTHNVTSESRVTKKAFATLSIAETFLFLYGVMDALIEKWFAS
 DRPGSIGISSILGLYLIGRAAFVFLSFLNLTAKAPNEKITWROOVYIWWAGLMR
 GAVSIALAYNFKTRSGHTQLHGNAMITSTIVVLFSTVMFGMMTPLIILLPLPASGH
 PVTSEPSPKSLHSPLLTSMQGSLESTINIVRPSLUMLLTKPTHTVHLYWRKFDDA
 LMRPFGGGRGVFPSPGSPTEQSHGR"
 2313
 polyA_site /note="17 a nucleotides"
 BASE COUNT 543 a 538 c 569 g 663 t
 ORIGIN

Query Match 35.3%; Score 768.8; DB 12; Length 2313;
 Best Local Similarity 68.6%; Pred. No. 1.2e-160;
 Matches 1077; Conservative 0; Mismatches 487; Indels 6; Gaps 1;

QY 312 gccttgatcagacatcgtcaagctctgtgtgctggtgaatctcttggtagcact 371
 DB 329 GGCTGTGTACACAGCTCCGATACCGCGTGGTGTGCTCAATCAACCTGTCGCGCGT 388
 QY 372 tcttggctgtgtattgtcttgggtcatcttttggaaagatagatgagacgaac 431
 DB 389 GCTCTGCGGCTGCATCTCTCTGGCACCTCTCGAGGAGATCGCTGGGCAATGAGTC 448
 QY 432 catcacgcctgtgtatggcttagcagcactgtgtgtaccatttggtagtagtaagg 491
 DB 449 CATCACGCGCTCATCATCGGCTCTGCACCGCGCTGGTGAATCTTCTCATCTACCTCCCTCC 508
 QY 492 aaaaagctgcacatctctctgttagtgaagatcttttctcatatctatttggcacc 551
 DB 509 GAAGAGCTGCACATTTATCTGCTCTCAGTGAGGATCTCTCTCATCTACCTCCCTCCCTCC 568
 QY 552 cattatattcaatgcaggttttbaagtaaaaagaagcagttttccgcaatttcgtgac 611
 DB 569 GATCATCTTCATGCAAGTTTLAGTAAAGAAAAGCAATTTCTCCGGAATTTTCATGAC 628
 QY 612 tattatgcttttggctgtgtggactattattcttgcacaaatcatatctctagggt 671
 DB 629 GATCACATTTATTTGGAGCGCTCGGCAATGATATCTCTTTTCAATATCTATTGCTGC 688
 QY 672 aacacagttcttaagaattgacattggaactttgacttgggtgattattctgcat 731
 DB 689 CATTGCAATTTACGACGAATGACATTTGGAACCGCTGGATGATGAGAGATTTCTTGCAAT 748
 QY 732 tgggtgcattatttggcgcaacagatttcagtatgtacactgcaggtttctgaatcaagada 791
 DB 749 TGGAGCCATCTTTTCTGCGACAGATTCCTGCTGCACATTTGCAAGTCCCTCAATCAGGATA 808
 QY 792 qacacattgctttacagctcttgattccgagagaggtgtgtgagatgatgcaagctcagt 851
 DB 809 GACACCTTTTGTGACAGTCTGATTTCTCGGTGAAGGTGTTTGAACGATGTATACATCAAT 868
 QY 852 tgtggttctcaacgcgattcagagctttgtatctcactcaactcaacacacgaagctgttt 911
 DB 869 TGTGCTTTTCAACGCCTACAGACATTTGATCTTGTCCACATAGATGCGGCTGTCGTCT 928
 QY 912 tcattctcttggaaactcttctgatttcttctcctaagtaacctgttgggtgctgcaac 971

Db 929 GAATTTTGGGAACCTCTCTTTTATTTTTCGAGCACCTTCCTTGGATATTTCG 988
 QY 972 cggctcgataagtcgctgtgttatacaagaactatactttgggaagccactcaactgacccg 1031
 Db 989 TGGATGTCTCAGTCATACATAATCAGAGCTATACATTTGGAGGCATTTACTACTACCG 1048
 QY 1032 agagtttcccttatgcttgaagctattcttcttcttcttcttcttcttcttcttcttcttct 1091
 Db 1049 TGAAGTGGCCCTTATGATGCTCATGCTTACCTTTATATATATATGTTGGCTGAGTTGTA 1108
 QY 1092 ctggagcgtatccctcactgtgtttctgtggtattgtgtgtccattacacabgca 1151
 Db 1109 TTTGAGCGCATCTCACCGTATTTCTCTGTGTTATTTGTAATTCACATACACTTGGCA 1168
 QY 1152 chatgaacgagactcaaaacaacaac 1211
 Db 1169 TAACTGCACAGAGATTTCAAGAGTTACAAACAGACGCAATTTGCAACTCTGCTTCAT 1228
 QY 1212 tgggagacattatttcttctgtgtgttgggaatgagctgtgacatttgacaagggag 1271
 Db 1229 TGTGAGACTTTCTCTCTCTGATTTGGGATGATGCAATTTGGATATTTGAAAAATGGGA 1288
 QY 1272 atccgtgagtaac 1331
 Db 1289 GTTGCAGTGACAGCTCTGGAATCCATTTGGATGATGCAATTTGGATGATTTGGATGATTTGGT 1348
 QY 1332 catgttgggaagcagcgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1391
 Db 1349 TCTGATTTGGAAGAGCTGCTTTTGTATTTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1408
 QY 1392 gaatacaagcagaagaatcaacttaacatcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1451
 Db 1409 GGCACCGAATGAAAATAATACCTGGAGACAGAAATGTAATATGTTGGTGGTGGTGGTGGTGGTGGT 1468
 QY 1452 gtaggtgctgtatctatgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1511
 Db 1469 GAGAGGAGCTGTGCTGATTTGCTTACATAAATTTTACAAGATCTGGCATACTCA 1528
 QY 1512 tatagcgggaatgaatcatcatcagcagatcagatcagatcagatcagatcagatcagatcagatcag 1571
 Db 1529 GCTGACGCGCAATGCAATTAATGATCACCAGCACCACCTGCTGCTACCGGCTCAGGCCATCC 1588
 QY 1572 ggtgttggtagtgac 1631
 Db 1589 GGTATTTGGGATGATGACAAAGCCATTTGATCAGGCTGCTGCTACCGGCTCAGGCCATCC 1648
 QY 1632 cagcagcagtttcttctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1691
 Db 1649 TGTACCC-----TCTGAGCTTTCATCACCAGTTCCTGCTGCTCCTCTCTCTCTCTCTCTCTCTCT 1702
 QY 1692 agactcgttcttggagcctttagggaac 1751
 Db 1703 CATGCAAGGTTCTGACCTCGAGTACAAACCAACATTTGTGAGGCTTCCAGCTCCGGAT 1762
 QY 1752 cttcttgacagcggcctcctggaacgctgattactgagacacacacacacacacacacacacacac 1811
 Db 1763 GCTCTCTCAACAGCGCCAC 1822
 QY 1812 catgcagcggccttcttggaggtcgt 1871
 Db 1823 GATGCGACCGATGTTTGGCGGCGGGGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1882
 QY 1872 gagaaacct 1881
 Db 1883 GCAGAGCCAT 1892

RESULT 6
 TM021B04/c
 LOCUS
 DEFINITION Arabidopsis thaliana BAC TM021B04.
 ACCESSION AF007271

PLN

DNA

30019 bp

TM021B04

12-JUN-1997

VERSION AF007271.1 GI:2191181

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

thale cress.

Arabidopsis thaliana

Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:

Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids II:

Brassicales: Brassicaceae: Arabidopsis.

1 (bases 1 to 90019)

Dante, M., Wamsley, P. and Gibson, A.

The sequence of A. thaliana TM021B04

Unpublished (1997)

2 (bases 1 to 90019)

Washington University Genome Sequencing Center.

The A. thaliana Genome Sequencing Project

Unpublished (1997)

3 (bases 1 to 90019)

Waterston, R.

Submitted (06-JUN-1997)

Submitted by:

Genome Sequencing Center

Department of Genetics, Washington University,

St. Louis, MO 63108, USA

e-mail: twilson@watson.wustl.edu

MAPING: Clones were assigned to the YAC map by hybridization by

M. Iohdi, Cold Spring Harbor Laboratories, and fingerprinted

by M. Marra, WashU, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone.

It may be shorter because we only sequence overlapping sections

once, or longer because we provide a small overlap between

neighboring submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded or sequenced with an alternate

chemistry; an attempt was made to resolve all sequencing problems,

such as compressions and repeats; all regions were covered by

sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The actual start of this clone is at base position 1 of TM021B04 ;

the actual end is at base position 90019 of TM021B04. The

orientation of this clone is unknown.

NOTES:

Coding sequences below are predicted from computer analysis, using

the program GeneFinder (P. Green and L. Hillier, ms in preparation).

Location/Qualifiers

1..90019

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="4"

/clone="TM021B04"

/map="5"

1802..3685

/gene="A_TM021B04.11"

join(1802..2050,2468..2686,2758..2844,2901..3095,

3375..3685)

/gene="A_TM021B04.11"

/note="contains similarity to synecococcus PCC7942

chromosomal region used as basis of neutral siteit

recombinational cloning vector (PID:g1174192)"

/codon_start=1

/evidence="not_experimental

/protein_id="AAB61092.1"

gene

CDS

1802..3685

join(1802..2050,2468..2686,2758..2844,2901..3095,

3375..3685)

/gene="A_TM021B04.11"

/note="contains similarity to synecococcus PCC7942

chromosomal region used as basis of neutral siteit

recombinational cloning vector (PID:g1174192)"

/codon_start=1

/evidence="not_experimental

/protein_id="AAB61092.1"

gene

CDS

1802..3685

join(1802..2050,2468..2686,2758..2844,2901..3095,

3375..3685)

/gene="A_TM021B04.11"

/note="contains similarity to synecococcus PCC7942

chromosomal region used as basis of neutral siteit

recombinational cloning vector (PID:g1174192)"

/codon_start=1

/evidence="not_experimental

/protein_id="AAB61092.1"

/db_xref="GI:2191197"

/translation="MVALHSVGLIPCSNQOKSFHSHYRNCIVCSSEFGLSIRRO

ALFVDSKSLSGDGERALSLVQDLQKPGKCFGAARQVQRLYTLEELKGLINAA

SLSPPTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

GLSGIVLDTTGGTIERNLQIAVSGGIVAKARGLDSQOFLTLGFMFLTLGVLEN

Db 65434 GTGGTAGAAGATCTCGATGATCAATGATTAATATTCTTATTGAATTGAC 65375
 Oy 1572 -----ggtgttgtagtgcgtaccacaa 1596
 Db 65374 ATTTGGTAATAGATTCCTTGAACCTGTTGTTGAGGTGTGATGACCAACCA 65315
 Oy 1597 ctcaaaagctacattacacagacagacagacagacagacagacagacagacagac 1656
 Db 65314 CTCATAGCTACCTATTACCGACAGAAACGACACAGACATGATCTGATGACAAAC 65255
 Oy 1657 accccaataac 1716
 Db 65254 ACCCAAAATCCATACATATCCCTTGTGACCAAGCTGTGATGACCTTGAAGG 65195
 Oy 1717 aaccacaatgtcctgcgcgtacagatagctgtgctcttgaacagccacacacac 1776
 Db 65194 AACCAATATGCTGCTGCGCTGACAGTATACGTGCTGCTGACACGCGCCACTGACAC 65135
 Oy 1777 gtagaatacctgagacacacacacacacacacacacacacacacacacacac 1836
 Db 134 GTGATTAATGATGACCAATTTGATGATCTTCTTATGACACCCGCTTGTGAGGTCGT 65075
 Oy 1837 ggtcttgtaacccctgttccaggttcccaactgagagaaacccctcctgaatgaag 1896
 Db 65074 GGCTTTGTACCTTTGTTCTCAGGTTCTCCACTGAGAGAAACCTCCTGATCTTAAG 65015
 Oy 1897 gcttgaaggtgaaggtgaagaaagcttgattttttttgtgtagaaaggtgattcaaa 1956
 Db 65014 GCTTGAAGGTGAAGGTGAAGAAAGCTTTGATTTTGTGTAAGAAAGGTGATTCAAA 64955
 Oy 1957 ttatgtctttgtgtaaatatcatcattgtaaatattgtttgtgagagacaaatcgtcct 2016
 Db 64954 TTATGCTTTGTGTAATATTCATTTGTAATATTTGTAAGACAGAAATCTGCTCT 64895
 Oy 2017 aacgtttgagagacaaagacaaacacacacacacacacacacacacacacac 2076
 Db 64894 AACGTTTGAAGACAGAAAGCAAAACATGCGCACTTGAAGTGTATGATGATGAT 64835
 Oy 2077 attatcatcattgtttgtttgttgaacacaaacacacacacacacacacacacac 2136
 Db 64834 ATTTATATTCATATTTGTTGTTGTAACAAACATTCATTTGTTGTTGATTT 64775
 Oy 2137 ggtttcgtcctgaa 2151
 Db 64774 GGTTTTGTCTTGA 64760

RI 7
 AB 16
 LOCUS AY028416 600 bp mRNA PLN 02-APR-2001
 DEFINITION Citrus x paradisi sodium/proton exchanger mRNA, partial cds.
 ACCESSION AY028416
 VERSION AY028416.1 GI:13908486
 SOURCE Citrus x paradisi.
 ORGANISM Citrus x paradisi
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 Sapindales; Rutaceae; Citrus.
 1 (bases 1 to 600)
 Portat.R., Laurie.S. and Povancello.D.
 A heat treatment induced the transcription of a sodium proton
 exchanger gene in grapefruit peel tissue
 Unpublished
 2 (bases 1 to 600)
 Portat.R., Laurie.S. and Povancello.D.
 Direct Submission
 Submitted (05-MAR-2001) Postharvest Sciences, ARO, The Volcani
 Center, Bet Dagan 50850, Israel
 Location/Qualifiers
 1..600
 /organism="Citrus x paradisi"
 /db_xref="taxon:37656"

CDS

/tissue_type="peel"
 /note="heat-induced"
 132..>600
 /note="ion transporter"
 /codon_start=1
 /product="sodium/proton exchanger"
 /protein_id="NAK27314.1"
 /db_xref="GI:13508487"
 /translation="MDQAISVVRKIQWVNSDHSVYSINIFVAPCASIVIGHLE
 ESRMNSITALLIGVAGVILLITGSGKSLFVSEDFEFTVLPPIFNAGROVK
 KKEFFNFTIMFGAIGLVGTCITISGIGQFFKRIIDIGYIAIGALF"
 BASE COUNT 170 a 86 c 141 g 203 t
 ORIGIN

Query Match 12/24; Score 266.6; DB 14; Length 600;
 Best Local Similarity 72/18; Pred. No. 4; A=49;
 Matches 347; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

Oy 265 tggatctagaagaagatbacaaatgttgatctctctagdytgcgaactcctctgacg 324
 Db 120 TGGTGCAGAAAATGGATCAAGCATATAGTTCTGTGTAAGAGAAATGCAATGGTGAAC 179
 Oy 325 aacatcgaacagcttctgtgtgctgtgaatcctctctgtgcaactcctgtgctgt 384
 Db 180 ACGTCTGATCAACAACGCGGTGCTTCCATTAACATTTTGTGGCCTTCCTGTGCTCT 239
 Oy 385 attgtcttgtaac 444
 Db 240 ATTGTGATTTGGTCTCTTCTTGAAGAAAGCGGTGATGACAGATTCATTCAGCTCTA 299
 Oy 445 ttgattggctaggaactggtgtaacacattgttgattagaagaagaagtcgcat 504
 Db 300 CTGATTTGAGATTTGTCAGAGGTGATATTTTCTGACAACTGGGGGAAACCTCACAT 359
 Oy 505 ctctcgtctttagaagaac 564
 Db 360 CTGTTTGCTTCAAGTGAAGATCTTCTTATATATGTTGTGCTCCCTCCATTTATAT 419
 Oy 565 gcaaggttcaagtaaaaaaagaagcatttccgaatttcgtgactatgaagcctttc 624
 Db 420 GCGATTTTCAGGTGAAAGAAAGCAATTTTCTCTAATCTTATCTATCATCATGCTG 479
 Oy 625 ggtgtgttggaactatattcttgcacaaatcattctcctgagtaacacacacac 684
 Db 480 GGTGCAATTTGGCACTATGTAAGCTGACATCATATCATAGGTGATTTAGTCTTT 539
 Oy 685 aagaagttggacattggaacacacacacacacacacacacacacacacacacacac 744
 Db 540 AAGAAATTTGATTTGTAATCTGATATGAGATTAATCTGCAATTTGTAATTT 599
 Oy 745 g 745
 Db 600 g 600

RESULT 8

AF279671 421 bp mRNA PLN 02-AUG-2000
 LOCUS AF279671
 DEFINITION Mesembryanthemum crystallinum putative Na⁺/H⁺-antiporter (HXH2)

ACCESSION AF279671
 VERSION AF279671.1 GI:9652169

KEYWORDS

common ice plant.
 Mesembryanthemum crystallinum
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
 Caryophyllales; Alceaceae; Mesembryanthemum.

REFERENCE 1 (bases 1 to 421)
 Chauhan,S., Forsthoefel,N. and Bohnett,H.J.
 Na⁺/myo-inositol symporters and Na⁺/H⁺-antiporters in plant sodium
 transport